UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS FÍSICAS

TESIS DOCTORAL

Lyman-alpha emitters in cosmological volumes

Emisores de Lyman-alpha en volúmenes cosmológicos

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Siddhartha Gurung López

Directores

Silvia Bonoli
Álvaro A. Orsi Moyano

Madrid

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS FÍSICAS
DEPARTAMENTO DE FÍSICA DE LA TIERRA
Y ASTROFÍSICA

TESIS DOCTORAL

Lyman-alpha emitters in cosmological volumes

Emisores de Lyman-alfa en volúmenes cosmológicos

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Siddhartha Gurung López

DIRECTORES

Silvia Bonoli
Álvaro A. Orsi Moyano

Madrid, 2019

Lyman-alpha emitters in cosmological volumes

Emisores de Lyman-alfa en volúmenes cosmológicos

Una memoria presentada por

Siddhartha Gurung López

y dirigida por

Dr. Silvia Bonoli
Dr. Álvaro A. Orsi Moyano

para aspirar al grado de Doctor en Astrofísica

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS FÍSICAS
DEPARTAMENTO DE FÍSICA DE LA TIERRA Y ASTROFÍSICA
Madrid, 2019

Lyman-alpha emitters in cosmological volumes

Emisores de Lyman-alfa en volúmenes cosmológicos

Siddhartha Gurung López

Universidad Complutense de Madrid
Facultad de Ciencias Físicas
Departamento de Física de la Tierra y Astrofísica

Centro de Estudios de Física del Cosmos de Aragón
Departamento de cosmología

UNIVERSIDAD
COMPLUTENSE
MADRID

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA LA OBTENCIÓN DEL TÍTULO DE DOCTOR

D./Dña. Siddhartha Gurung López
con número de DNI/NIE/Pasaporte 52898418M, estudiante en el Programa de Doctorado de Astrofísica,
de la Facultad de Ciencias Físicas de la Universidad Complutense de Madrid, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

Lyman-alpha emitters in cosmological volumes // Emisores de Lyman-alfa en volúmeres cosmológicos.

y dirigida por: Silvia Bonoli y Álvaro A. Orsi Moyano

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 30 de abril de 2019

Fdo.:

Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

“ 1.048596 ”

Steins;Gate

Para tí, madre

and

to whoever was, is and will be.

Dedicación

Esta tesis y el trabajo que se le ha dedicado están dedicados a todas las personas que me han ayudado y enseñado a lo largo de mi vida.

Me gustaría empezar por mi familia. A toda. Y en particular para las personas que tanto me han ayudado durante estos 4 años: La mas mejor de todas las tias; mi tia sabel, que me ha cuidado toda mi vida. Mi prima Ruth y su pequeñaja Tábata. A mis primos Ángel y Gerson. Sin ellos, esta tesis no hubiera sido posible. Viajes como el de a Disney me ayudaron a seguir a delante.

Para mis amigos voy a ir por order cronológico. Por ello empiezo con Miguel, Prince y Borja. Con nuestros mas y nuestros menos, hemos sido amigos durante el 89% de nuestras vidas actuales, ya que nos conocimos a los 3 años. Fuimos juntos a la guardería, al colegio, al instituto. Increíbles personas.

Prácticamente al mismo tiempo conocí a mis amigos de Las islas. Juan, Nacho, Juanito, Gonzalo, Ángel, Álvaro, Nachito, Carlitos. Mas tarde se nos uniría Victor. Sin lugar a dudas, los veranos que pase en casa de mi tia en Las islas fueron de los momentos mas felices de mi vida. Y todo esto, gracias a mi amigos. Piscina, pachanga, piscina, pachanga. No ha habido ni habrá una mejor rutina. He dicho. Claro está, mas tarde se nos unirían Ro, Bita, Roxana y otros muchos durante nuestra adolescencia y sus primaverales consecuencias. En la actualidad nos encontramos todos un poco esparcidos. Tal vez algún día todo vivamos de nuevo en el mismo Pais y podamos recuperar la tradición del Culo durante el día y el Pro durante la noche.

De Las islas, no solo quiero agradecer a la gente de mi edad. También a los padres de mis amigos, que han sido como una familia para mí desde una tierna edad hasta los momentos mas difíciles. A todos, pero en particular a Josefa y Juvenal, a Ester y

Carlos, a Mari Carmen y José Antonio.

Llegamos a mi época en el instituto. He de reconocer que no llegue como buen estudiante, pero me gusta pensar que salí siéndolo. Esto fue gracias a mi madre, a mis amigos y por su puesto, a los profesores que me dieron clase en el Gómez Moreno. No solo consiguieron que me memorizase y entendiera lo necesario para pasar de curso. Consiguieron algo más importante, que tuviera curiosidad por el conocimiento, cosa de la que carecía hasta los $\sim$15 años. Creo que mencionar esto en una tesis doctoral es de extrema importancia. En ciencias Ana Barberia, Maria José, Rosa y Pagola hicieron el trabajo. En letras Ana Matei y Cristobal. Cabe destacar, que sin compañeros como Laura Ramos, Ferrero, Bernet, Delgado, Amanda, Alex, Piña, Javi, Sara, Ivan, Ya Zhen las clases hubieran sido mucho un poquito aburridas y nada hubiera sido lo mismo.

Llegue, de alguna manera, a la universidad. Donde conocí a gente increíble. En primero me cruce con Beah. En segundo, nos juntamos prácticamente todos los que quedábamos por juntarnos. Fue entonces cuando cuando conocí al holgazán de Garcés (que ahora trabaja más que nunca), al gañan de David Gómez, al Guantes, a Gabixuela y Paloma, a Miguel y finalmente al mayor fraude que he conocido y que jamás conoceré: Miguel Ángel. Mas tarde se nos uniría la bellísima persona de Maria Angeles. Durante estos años muchas cosas pasaron, muchas columnas, muchas (una, pero aún así demasiadas) noches siendo perseguidos por pueblos Murcianos, muchísimas clases divertidas, muchísimos momentos inolvidables, muchísimos Fridays, Fosters, Ginos. Inolvidable, y que por tanto no es posible de olvidar fue el viaje a Canarias. Espero que podamos seguir pasando tiempo juntos en lo que nos quede.

Por su puesto los profesores de la facultad de física de la UCM también han jugado un papel muy importante. Ellos continuaron con la labor de mis profesores de instituto. Cabe remarcar a Artemio. El mejor profesor que he tenido y una bellísima persona. Una fuente de inspiración. Pero otros profesores como Maroto, Martín Delgado, Carmelo, Gorgas, Gallego, Armando no se quedan a tras. Una gran parte de los profesores que me dieron clase en la Universidad me inspiraron para seguir aprendiendo e investigar los misterios del cosmos.

Y llego Teruel, y con Teruel llego CEFCA, donde he pasado los últimos 4 años de mi vida. Es sabido que el doctorado no es una etapa fácil. Sin embargo toda la gente que he conocido en Teruel, y en particular en el corredor donde ahora mismo me encuentro escribiendo, me han ayudado a seguir a delante. Todos me han enseñado algo. Pero creo que Gonzalo se merece un reconocimiento especial por enseñarme a usar python. No que hubiera hecho sin los italianos. Daniele por ser una gran persona, las noches de películas y Flanagans. Mateo por su humor y por la pasta a las 4 de la mañana siempre que era necesaria. Gio por ser Gio. Y en particular por los míticos partidos de futbol, en los que por su puesto, también participaba más gente. Pero raramente 10. A David y sus ganas continuas de fiesta. A Rafa y sus maravillosa manera de despotricar. A Angela y Cecile. A Jonas y al resto de las personas que han hecho de CEFCA y Teruel un hogar para mi estos 4 años.

De CEFCA cabe destacar, por supuesto a mis directores de tesis. A los dos. A Álvaro y a Silvia. Que más que supervisores, han sido amigos. Ahora soy consciente de lo regulinchi que era como estudiante al principio. La paciencia que Álvaro y Silvia han tenido conmigo no ha sido precisamente poca. Y se lo agradezco. Les agradezco muchísimo todo lo que me han enseñado durante estos años. Y no solo incluyo conocimiento, sino en maneras de pensar y tener una visión científica critica. Desde mi sesgado punto de vista, son los mejores directores que un estudiante de doctorado puede tener. Gracias.

Acknowledgements

I would like to thank to Masai Ouchi for hosting me in his team for two moths. I really enjoyed my time in Japan and learnt a lot thanks to him and his team.

Also, I would like to thank Nelson Padilla for hosting me another two months that made this thesis go further.

The authors acknowledge the support of the Spanish Ministerio de Economía y Competividad project No. AYA2015-66211-C2-P-2. We acknowledge also STFC Consolidated Grants ST/L00075X/1 and ST/P000451/1 at Durham University. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

Abstract

This thesis studies how the complex radiative transfer processes that Lyman-α photons suffer within atomic hydrogen clouds shape the observability of galaxy samples selected by their Lyman-α flux.

First, we make a theoretical introduction to Lyman-α radiative transfer processes. The quantum description of the hydrogen atom predicts that the hydrogen atom exhibits a discrete structure of energy levels, which explains the different emission lines found in the hydrogen spectrum, and in particular Lyman-α. Quantum mechanics also predict that the probability of interaction between a Lyman-α photon and a hydrogen atom is very high. This causes that photons entering into a cloud of atomic hydrogen are absorbed and reemitted constantly, changing their frequency and direction in each scattering event. In this way, the length of the trajectory of Lyman-α photons is increased, which makes them very sensitive to destruction mechanisms, such as dust.

The properties of the Lyman-α radiation that emerges from neutral hydrogen clouds depend greatly on the properties of the gas. By making use of a galaxy formation model run over a large N-body simulation, we find that the gas geometry, its macroscopic velocity with respect to the Lyman-α source, the atomic hydrogen density and the dust content are critical. The fraction of photons that escape from the gas is found to be heavily determined by its metallicity, although other properties also play a significant role. The line profile, i.e., the frequency distribution of photons, is also influenced by all these properties. The reprocessing of the line profile is such that initially monochromatic Lyman-α radiation might emerge exhibiting several emission peaks. In general, the higher is the number of interactions with neutral hydrogen, the more Doppler shifted are the Lyman-α line profiles.

Lyman-α radiative transfer processes have a direct impact in the selection function of galaxy samples detected by their Lyman-α flux. Inside galaxies, this radiation is generated in the hot surroundings of very young stars. Then, the Lyman-α photons

enter into the cold and dusty interstellar medium partially composed by hydrogen. There, Lyman-α radiation gets easily attenuated by dust grains. This causes that Lyman-α selected samples are mainly constituted by metal poor star forming galaxies of intermediate mass.

After escaping the original galaxy, the remaining Lyman-α photons enter into the cold intergalactic medium, mainly populated by atomic hydrogen. There, the radiative transfer processes continue. We find that the selection function of Lyman-α selected samples depends on the large scale properties of the intergalactic medium. In particular, environments with lower content of neutral hydrogen are predicted to be more transparent to this radiation. Additionally, if the emitting galaxy and the gas along the line of sight are moving away from each other, the Lyman-α photons received by the atomic hydrogen are Doppler shifted to reader wavelengths. This decreases the number of scattering events and enhance the transmission. In this way, galaxies in our models with Lyman-α radiation detected tend to lie in low dense environments in which it is more probable that the galaxy is distancing from the surrounding intergalactic medium. This has a direct impact in the measured clustering of this galaxy population. This is of crucial importance for future surveys targeting this population of galaxies to study cosmology.

At some point during the evolution of the Universe, the first stars and galaxies were born surrounded by a neutral and cold intergalactic medium. The energetic radiation escaping from galaxies is expected to create small bubbles of ionized gas around them. The size of these bubbles grows until they overlap and cover the whole Universe. This is known as the epoch of reionization of the Universe. During this time window, the coupling between the intergalactic medium and its transmission to Lyman-α radiation is predicted to be very strong. At first order, while Lyman-α radiation could not escape form neutral regions, ionized regions were transparent. In this way, the distribution of galaxies measured with Lyman-α emission is predicted to track the topology of neutral and ionized hydrogen in the Universe. As a consequence, the clustering of galaxies selected by their Lyman-α flux at this epoch is a great tool to constrain the history of reionization of the Universe.

Finally, we present the first results of HiRULE (High Redshift Universe Lyman-α survEy). The main goal of this survey is to study the cross-correlation between galaxies which Lyman-α emission comes from star formation episodes and quasars. For this,

we scan the sky looking for galaxies at redshift ~ 2.2 with Lyman-α emission. For this goal, we make use of narrow and broad photometric bands. The broad bands are used to detect the continuum of distant galaxies while the narrow bands allow us to measure the Lyman-α flux of our targets. Despite HIRULE being in an early phase of development, the results we find are promising. In particular, the abundance of LAEs in our survey matches the one observed by other scientific teams.

Resumen de la tesis

Esta tesis estudia cómo los complejos procesos de transferencia radiativa que sufren los fotones Lyman-α dentro de nubes de hidrogeno neutro determinan la function de selección de muestras de galaxies detectadas por su flujo en Lyman-α.

Primero, hacemos una introducción teórica a los procesos radiativos de Lyman-α. Es de gran importancia la descripción cuántica del átomo de hidrogeno. Este formalismo predice que el átomo de hidrogeno presenta una estructura discreta de nivele de energía, el cual explica las diferentes lineas de emisión encontradas en el espectro del hidrógeno, y en particular Lyman-α. La mecánica cuántica también predice que la probabilidad de interacción entre un fotón Lyman-α y un átomo de hidrógeno es muy elevada. Esto causa que cuando uno de estos fotones entra en una nube de hidrogeno neutro, este es absorbido y reemitido constantemente, cambiando su frecuencia y dirección en cada interacción. De esta manera, la longitud de la trayectoria de los fotones Lyman-α se incrementa, lo cual hace esta radiación muy sensible a mecanismos de destrucción, como el polvo.

Las propiedades de la radiación Lyman-α que emerge de nubes de hidrógeno neutro dependen en gran medida de las propiedades del gas. En particular, la geometría del gas, su velocidad macroscópica con respecto a la fuente de Lyman-α, la densidad de hidrógeno neutro y el contenido de polvo son críticos. La fracción de fotones que escapan del gas esta altamente determinado por su metalicidad, aunque las otras propiedades también tienen su papel. Adicionalmente, el perfil de linea, es decir, la distribución de frecuencias de los fotones, esta muy influenciada por todas estas propiedades. El reprocesamiento del perfil de linea es tal que, radiación Lyman-α inicialmente monocromática puede emerger con varios picos de emisión. En general, cuanto mayor es el numero de interacciones con hidrógeno neutro, más desplazadas están las lineas de Lyman-α por efecto Doppler.

La transferencia radiativa de fotones Lyman-α tiene un impacto directo en la función de selección de muestras de galaxias detectadas por su flujo en Lyman-α. Dentro de las galaxias, esta radiación es generada in los calientes alrededores de estrellas jóvenes. Inmediatamente después, los fotones Lyman-α entran en el frío medio interestelar rico en polvo, parcialmente compuesto de hidrogeno, donde la radiación Lyman-α es fácilmente atenuada por granos de polvo. Esto causa que muestras seleccionadas por Lyman-α están principalmente formadas por galaxias con baja metalicidad y de masa intermedia.

Después de escapar de su galaxia original, los fotones Lyman-α que quedan entran en el frío medio intergaláctico, poblado principalmente por hidrógeno atómico, donde los procesos de transferencia radiativa continuan. Esto causa que la función de selección de muestras basadas en la detección de Lyman-α dependa de las propiedades a gran escala del medio intergaláctico. En particular, los entornos con menor densidad de hidrógeno neutro son más transparentes a esta radiación. Adicionalmente, si la galaxia emisora y el gas a lo largo de la linea de visión se distancian el uno del otro, los fotones Lyman-α son recibidos por los átomos de hidrógeno neutro corridos a longitudes de onda mas rojas. Esto baja el numero interacciones y aumenta la transmisión. De esta manera, galaxias en las que se detecta radiación Lyman-α tienden a habitar en entornos de baja densidad en las que es mas probable que la galaxia se este alejando del medio intergaláctico que las rodea. Esto tiene un impacto directo in la distribución espacial de esta población de galaxias. Esto es de crucial importancia para futuros cartografiados astronómicos que usaran este tipo de galaxias para estudiar la cosmología del Universo.

En algún momento durante la evolución del Universo, las primeras estrellas y galaxias nacieron rodeadas de un frío y neutro medio intergaláctico. La radiación energética que escapaba de las galaxias creo pequeñas burbujas de gas ionizado al rededor de estas. Poco a poco el tamaño de estas burbujas fue creciendo, hasta que estas se empezando a solapar y cubrieron todo el Universo. Este lapso de tiempo es conocido como la época de reionización del Universo. Durante este periodo, el acoplamiento entre el medio intergaláctico y su transmisión a radiación Lyman-α se vuelve muy fuerte. En primera aproximación, mientras que la radiación Lyman-α no puede escapar de las regiones neutras, las regiones ionizadas son transparentes. De esta manera, la distribución de galaxias detectadas con emisión Lyman-α se asemeja a la topología del hidrógeno neutro e ionizado en el Universo. Como consecuencia, la disposición en el cielo de galaxias seleccionadas por su flujo en Lyman-α durante esta época es una gran herramienta para determinar la historia de reionización del Universo.

Finalmente presentamos los primeros resultados de HiRULE (High Redshift Universe Lyman-α survEy). El principal objetivo de este cartografiado es estudiar la correlación cruzada entre galaxias que emiten Lyman-α debido a su formación estelar y cuasares. Para ello, escaneamos el cielo en busca de galaxias con un desplazamiento al rojo ~ 2.2 con emisión en Lyman-α. Con este objetivo, hacemos uso de bandas fotométricas anchas y estrechas. Las bandas anchas se usan para detectar el continuo de galaxias lejanas, mientras que las bandas estrechas nos permiten medir el flujo en Lyman-α de nuestros candidatos. HiRULE se encuentra en una fase muy temprana de desarrollo, pero hasta ahora hemos obtenido resultados prometedores. En particular, la abundancia de este tipo de galaxia en nuestro cartografiado concuerda con la observada en otros estudios científicos.

Table of contents

List of figures

List of tables

List of acronyms

ΛCDM	Λ Cold Dark Matter.
2PCF	2-Point Correlation Function.
AGN	Active Galactic Nuclei.
BAO	Baryon Acoustic Oscillation.
BB	Broad Band.
CMB	Cosmic Microwave Background.
DM	Dark Matter.
EW	Equivalent Width.
FWHM	Full Width Half Maximum.
HOD	Halo Occupation Distribution.
IGM	Intergalactic Medium.
IMF	Initial Mass Function.
ISM	Interstellar Medium.
KS	Kolmogorov-Smirnov.
LAE	Lyman-α Emitter.
LF	Luminosity Function.
LoS	Line of Sight.
Lyα	Lyman-α.

MCMC	Monte Carlo Markov Chain.
MCRT	Monte Carlo Radiative Transfer.
NB	Narrow Band.
QSO	Quasi Stellar Object.
RT	Radiative Transfer.
SED	spectral energy distribution.
SFR	Star Formation Rate.
sSFR	specific Star Formation Rate.

Chapter 1

Introduction

This thesis verses about Lyman-α emission, the physical processes affecting this radiation and the galaxy population known as Lyman-α emitters (LAEs). We dedicate this chapter to give a brief description of the current standard cosmological model, a theoretical overview to the complex Lyα radiative transfer processes and a historical introduction to this interesting galaxy population.

1.1 Current cosmological paradigm

In this section we give a compact introduction to the current standard cosmological model. Most of the material of this section comes from Mo et al. (2010), Lyth and Liddle (2009) and Amendola and Tsujikawa (2010).

The current standard cosmological model is ΛCDM. One of the main pillars of ΛCDM is the *cosmological principle*, which states that the Universe is homogeneous and isotropic on scales large enough. The cosmological principle allows the diversity of *small* structures that we observe in the Universe, such as planets, stars, galaxies, filaments and voids among others. However, it provides a physical scenario where, statistically, the properties of these *small* structures in a large enough Universe volume is the same independently of the piece of the Universe that was observed.

ΛCDM gives a description for the evolution of the Universe after its birth, the *Big Bang*. We summarize the evolution of the Universe in Fig. 1.1. The standard cosmological model starts the narrative of the life of the Universe when all its energy was encapsulated in a single point. The time window when the Universe was younger than

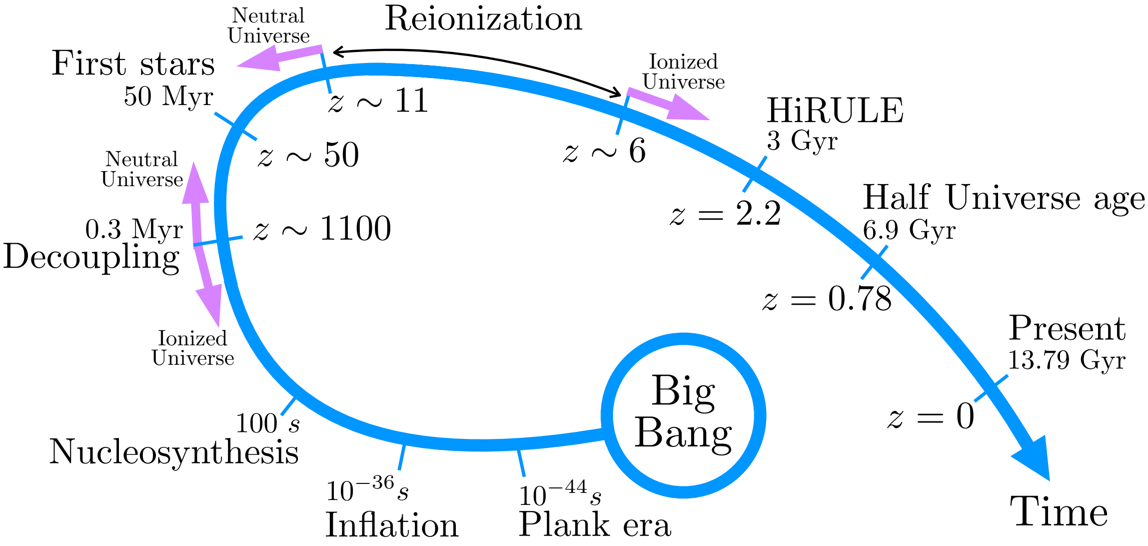


Fig. 1.1 Timeline of the history of the Universe. The blue arrow indicates the time flow. We indicate an approximate redshift and age of the Universe for each event. The age was computed assuming Planck Collaboration et al. (2016) cosmology. The purple arrows indicate the moments when the Universe becomes ionized or neutral. We also show the time window of HiRULE, a galaxy survey that we describe in Chapter 6.

the Plank time ($\sim 10^{-44}s$) is known as the Plank epoch. During this time our current knowledge about physics is not enough to understand the physical processes taking place. This era remain, up to date, one of the most mysterious physical environments that humankind has ever conceived.

At some point after the Plank era and before the formation of the first protons and electrons, *inflation* took place. Inflation suggested in first place by Guth (1981). One of motivations behind inflation is that the Big Bang theory predicts that the Universe was causally connected shortly after the Big Bang and in thermal equilibrium. Therefore everything was smooth and homogeneous. However, today we observed that the Universe exhibits a large scale structure with high density contrasts. Another of the motivation was the extreme flatness of the Universe. In short, inflation predicts that in a very early epoch of the Universe ($\sim 10^{-36}$ s) there were quantum density perturbation in the Universe. Then, for a short period of time, the expansion of the Universe became accelerated in such a way that the size of the quantum perturbations became comparable with the size of the causally connected Universe. In this way, after inflation, there are different cosmic regions with high density contrast. This contrast does not vanish after inflation as the regions are causally disconnected. Additionally, inflation also solves the flatness problem.

In the Big Bang theoretical framework, the Universe has been expanding since its birth. As the Universe expands it gets progressively colder. Eventually, when the Universe was close to be one second old, it reached the temperature low enough to form baryons such us protons and neutrons. Shortly after this, when the Universe becomes slight colder, leptons such as electrons and neutrinos are formed.

At this stage, neutrinos, electrons, protons, neutrons and photons were at thermal equilibrium. In other words, the mean free path of these particles was significantly smaller than the size of the causally connected Universe. However, when the temperature drops below $\sim 1\ MeV$, the interaction rate for neutrinos falls and the mean free path becomes larger than the observable Universe. This is known as *neutrino decoupling*. Indeed, when neutrino decoupled and were able to travel freely, they still had a relativistic temperature. This causes the *neutrino background*.

Shortly after this, when the Universe was $\sim 100\ s$ old, protons and neutrons started to bind together to form light nuclei. This is known as *nucleosynthesis*. The relative abundance of the synthesized nuclei. ΛCDM predicts that $\sim 74\%$ of the protons remain by their own and constituted the nuclei of the hydrogen atoms formed later (^{1}H). About a 25% of the formed nuclei were ^{4}He, as it is the light nuclei with the biggest binding energy. The rest of the synthesized nuclei were ^{2}H, ^{3}He and ^{7}Li.

Until this point, photons frequently interact with free electrons through Thomson scattering. In this way, photons and matter were coupled and in thermal equilibrium. However, as the Universe keeps expanding, the temperature drops lower, allowing nuclei and electrons to bind together. This is known as *recombination*. After recombination, the number of free electrons decreased dramatically and matter and photons became decoupled. The frequency distribution of the decoupled photons was as of a black body with temperature equal to the temperature of the thermal equilibrium just before decoupling. This radiation is known as the *cosmic microwave background* (CMB).

Let us define now the *geometric redshift* z_g. In general, any two points, between which the Universe is expanding, are distancing. Thus, if we consider two sources without proper motion lying, each of them, in one of those points, each of the particles sees the other moving radially away from them. Therefore, the radiation received by each source is Doppler shifted towards redder wavelengths, i.e., redshifted. In astrophysics and cosmology the geometric redshift is defined as the Doppler shift due to the expansion of the Universe measured at Earth in radiation emitted at a certain epoch, i.e.,

$$z_g = \frac{\lambda_{\mathrm{g,obs}}}{\lambda_{\mathrm{emi}}} - 1, \tag{1.1}$$

where λ_{emi} is the wavelength of the photons emitted in the source frame and λ_{obs} is the wavelength measured by the observer due to the expansion of the Universe.

In practice, the measured redshift z of galaxies and other objects is not purely the geometric redshifts. Indeed, there are additional components, such us the peculiar motion of the sources, that modifies the observed redshift. These other contributions are typically smaller than 10^{-2}. Although small, they have an impact in the apparent distribution of galaxies in the sky (Kaiser, 1987).

ΛCDM predicts that the energy of the Universe is divided into four components: radiation, matter, the expansion of the Universe and its curvature. The weight of each of each components determines the expansion rate of the Universe, in particular,

$$H^2(z) = H_0^2 \left[\Omega_r(1+z_g)^4 + \Omega_m(1+z_g)^3 + \Omega_k(1+z_g)^2 + \Omega_\Lambda\right], \tag{1.2}$$

where Ω_r, Ω_m, Ω_k and Ω_Λ are the energy density of radiation, matter, curvature and expansion of the Universe respectively. Meanwhile, H_0 is the Hubble parameter at redshift 0. The Plank experiment (Planck Collaboration et al., 2016) and other experiments such as WMAP (Wilkinson Microwave Anisotropy Probe, Bennett et al. (2013)) have estimated the value of the different energy densities by measuring the CMB spectrum. In particular, Plank computed that our Universe has $\Omega_r \sim 5 \times 10^{-5}$, $\Omega_m = 0.305$, $\Omega_\Lambda = 0.691$. Additionally they also found that our Universe extremely flat, thus $\Omega_k \sim 0$.

The other two pillars of ΛCDM a part from the cosmological principle and inflation are i) the presence of cold dark matter in addition the regular matter and ii) the assumption that Ω_Λ is constant in time. On one hand, the first assumption predicts that a part from the ordinary matter composed by neutrinos, electrons, protons and so on, there is a matter component that a) with negligible interaction with itself, light or any other thing (dark) and b) has negligible random motion (cold). On the other hand, the assumption that Ω_Λ does not evolve is also known as *dark energy.* These two pillars are still to be confirmed or excluded by the scientific community. For example, it has not been found yet a particle with the characteristics of the dark matter. Neither the scientific community has been able to measure with enough precision Ω_Λ through cosmic time. However, the number of observables that ΛCDM is able to reproduce, such us the accelerate expansion of the Universe and the velocity radial profile of galaxies, is overwhelming Mo et al. (2010).

After decoupling the Universe was in a neutral state. At this epoch the Universe was still quite smooth in comparison with nowadays. This epoch is known as the *dark ages*. During this time gravity becomes dominating and the formation of structures begin. This is known as the hierarchical growth of structures. The regions with the highest density contrast accrete more material and become denser little by little. In the regions with the highest density contrast the first dark matter halos start to form and to be virialized. Inside the dark matter halos, the baryonic gas (mostly atomic hydrogen and helium) colds through several mechanisms, such us radiative cooling, Bremsstrahlungn cooling and Compton cooling. Eventually, the atomic hydrogen is cold enough to for molecular hydrogen and consecutively form the first stars and galaxies.

The birth of the first stars and galaxies put and end dark ages and marked the begging of a Universe flooded with bright sources. Inside galaxies there are episodes of star formation, where cold gas collapses, fragmented, keeps collapsing and form stars with different masses. The spectrum of the radiation emitted by stars is strongly linked to their mass. The heavier a star is the more energetic are the photons emitted. Indeed, the so called O-type and B-type stars started injecting photons capable of ionizing atomic hydrogen intro the Universe. In this way, the regions surrounding galaxies became ionized forming bubbles. Eventually, the bubbles around each galaxy started to overlap with each other until the Universe was completely ionized. This time window is known as the epoch of reionization (Barkana and Loeb, 2001).

As the reader might have noticed by now, there is a lot of *darkness* in the current standard cosmological model, but, fortunately, there is also a lot of *light*. Nowadays, one of the most spread methodologies to study cosmology is to scan the sky looking for galaxies. Galaxies are formed in the regions with the highest density and therefore, they trace the underlying density field of dark matter and baryons. This is very useful, as the properties of the density fluctuations are a direct prediction of ΛCDM.

One of the main observational evidences of the dark energy is the baryon acoustic oscillation (BAO), which corresponds to a clear imprint in the clustering of galaxies on large scales. The BAO were created early in the Universe, when baryons and photon were tightly coupled, before recombination. During this epoch, the oscillation produced by sound waves modifies the density field, which translates into a modification in the spatial distribution of galaxies. In practice, the BAO creates a peak in the 2-point correlation function of galaxies around ~ 100 cMpc/h.

The scientific community is making a huge effort to unveil the nature of dark energy. With this goal in mind, the BOSS galaxy survey (Alam et al., 2017) has already measured the BAO. Additionally, many galaxy surveys, such us J-PAS (Benitez et al., 2014a), HETDEX (Hill et al., 2008a), EUCLID (Moresco et al., 2018), DESI (DESI Collaboration et al., 2016) are planning to measure the BAO in larges redshift windows and with better precision. From these surveys, J-PAS and HETDEX are of special importance for this thesis, as they plan to trace the density field using a galaxy population known as Lyman-α emitters. This galaxy population is observed by their strong emission of Lyα photons. However, Lyα is a resonant line that undergoes important radiative transfer processes inside clouds of atomic hydrogen. Some authors, such us Zheng et al. (2011), have explore the impact of the Lyα radiative transfer on the selection function of Lyα emitters and how this might affect their measured clustering. Any clustering modifications might translate into systematic errors if the upcoming data in not well analyzed. Therefore, it becomes crucial to understand properly this galaxy population.

1.2 The hydrogen atom

In this section we introduce the hydrogen atom and its classical and quantum descriptions. We study closely the mechanisms of Lyman α photon emission and their cross-section with neutral hydrogen atoms. Most of the material used in this section comes from Griffiths (1982) and Dijkstra (2017).

1.2.1 Classical description

Hydrogen is the lightest element of the periodic table and it is composed by a proton and an electron. In classic mechanics the proton is static while the electron orbits around it (see left panel of Fig.1.2). In particular, the centrifugal force experienced by the electron is compensated by the Coulomb attraction to the proton, i.e.,

$$m_e \frac{v_e^2}{r} - \frac{1}{4\pi\epsilon_0} \frac{q^2}{r^2} = 0 \tag{1.3}$$

where m_e is the electron mass, v_e is its velocity, r is the distance to the proton, ϵ_0 is the vacuum permittivity and and q is the absolute value of the electron charge. Additionally, the total energy of the electron is simply the sum of kinetic and potential energy

$$E_T = \frac{1}{2} m_e v_e^2 - \frac{1}{4\pi\epsilon_0} \frac{q^2}{r} = -\frac{1}{4\pi\epsilon_0} \frac{q^2}{2r}. \tag{1.4}$$

This system is unstable due to the fact that accelerated charged particles emit radiation. In particular, the power (energy per unit of time) radiated by the electron is given by the Larmor formula:

$$P = \frac{2}{3} \frac{q^2 a_e^2}{4\pi\epsilon_0 c^3} \tag{1.5}$$

where a_e is the modulus of the acceleration experienced by the electron and c is the speed of light in vacuum. Combining Eq.1.3 and 1.5, the necessary time for the electron to radiate all its energy is

$$t = \frac{E_T}{P} = (4\pi\epsilon_0 m_e)^2 \frac{3r^3c^3}{4q^4} \tag{1.6}$$

where if r is substituted by the Bohr radius[1] ($5.3\ 10^{-11} m$), the typical decay time is $\sim 10^{-11} s$. In this way, the classical version of the hydrogen atom is clearly unfeasible, since, in reality, we find hydrogen in nature. In fact hydrogen nuclei constituted the 74% of the total number of nuclei formed during the Big Bang nucleosynthesis.

1.2.2 Quantum description

The short life time of the hydrogen atom in its classic version and other, by the time inexplicable, physical phenomena such us the black body radiation (discovered by Kirchhoff in 1860) lead to the development of quantum mechanics. In the quantum interpretation of the hydrogen atom (see right panel of Fig.1.2) the electron inhabits discrete states with discrete energy levels. Despite the classic picture where the electron

---

[1] The Bohr radius is defined as $r_B = \epsilon_0\ h^2 / m_e q^2 \pi$.

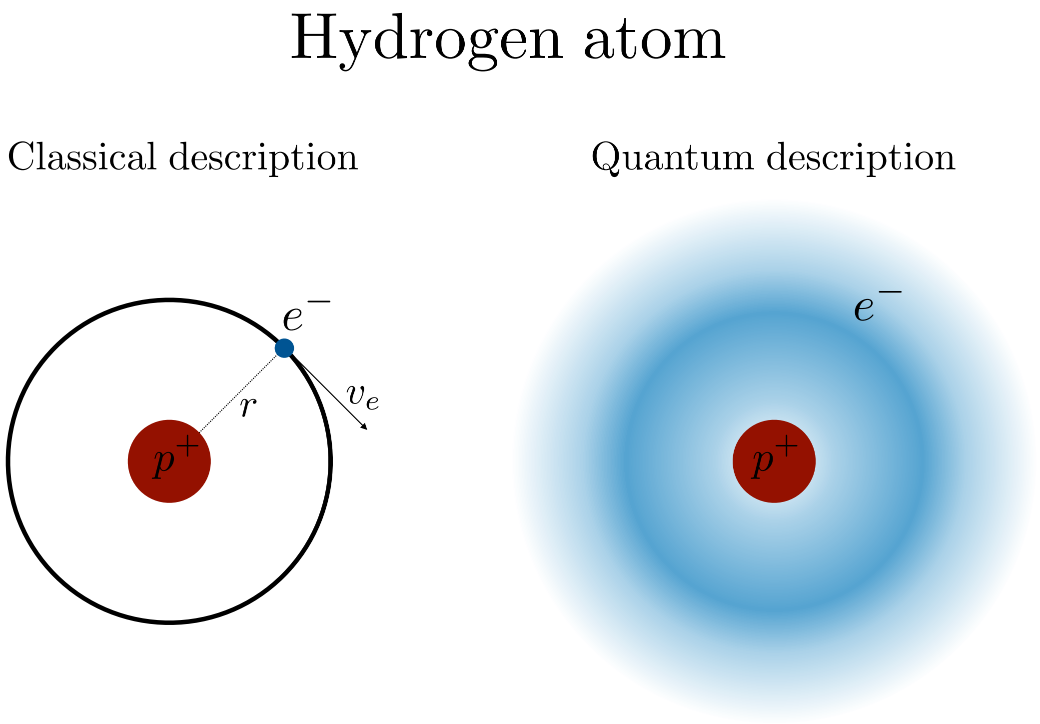


Fig. 1.2 Artistic sketch of the classic (left) and quantum (right) descriptions of the hydrogen atom. While in the classical version the electron is a particle with a well define trajectory, in the quantum version it inhabits a cloud of probability.

follows a well defined trajectory, in quantum mechanics the electron lies in a probability cloud, free of acceleration.

The hydrogen atom constitutes one of the few systems with analytic solution for the Schrödinger equation. The time independent Schrödinger equation for a non-relativistic particle with potential energy $V(r)$ is written as

$$E\Psi(\mathbf{r}) = \left[\frac{-\hbar}{2\mu}\nabla^2 + V(r)\right]\Psi(\mathbf{r}) \tag{1.7}$$

where $\Psi(\mathbf{r})$ is the wave function describing the stationary state of a particle giving the probability of finding it in the location $\mathbf{r}$, E is the energy of the particle in the state $\Psi(\mathbf{r})$, $\hbar$ is the reduced Plank constant, μ is the reduced mass of the particle and ∇^2 is the Laplace operator. In particular, in the hydrogen atom, the potential energy of the electron is

$$V(r) = -\frac{1}{4\pi\epsilon_0}\frac{q^2}{r}. \tag{1.8}$$

It can be demonstrated (Griffiths, 1982) that in the hydrogen atom, the electron wave function $\Psi(\mathbf{r})$ can be described by four quantum numbers. These are:

1. The principal quantum number n ($n \geq 1$), which determines the total energy of the electron
$$E_n = \frac{E_1}{n^2}, \tag{1.9}$$
where $E_1 = -13.6\text{eV}$ is the energy of the ground level.

2. The orbital quantum number ℓ, which gives the angular momentum
$$L = \hbar\sqrt{\ell(\ell+1)}, \tag{1.10}$$
where ℓ is integer from 0 to $n-1$ and $\hbar$ is the reduced Plank constant.

3. The magnetic quantum number m_ℓ, which provides the direction of the electron angular momentum
$$L_z = m_\ell \hbar, \tag{1.11}$$
where m_ℓ takes integer values between $-l$ and l.

4. The spin projection quantum number m_s, which escribes the intrinsic angular momentum of the electron within an orbital along one axis
$$S_z = m_s \hbar, \tag{1.12}$$
where $m_s = -s, -s+1, ..., s-1, s$ with s the intrinsic spin of the particle ($s = 1/2$ for electrons).

Historically, a state with quantum numbers $\{n,\ell\}$ is called by the value of the its principle quantum number and the letters "s", "p", "d", "f",... when $\ell = 0, 1, 2, 3, ...$. In this way, the state $\{n = 1,\ \ell = 0\}$ is called "1s", the state $\{n = 2,\ \ell = 0\}$ is called "2s", the state $\{n = 2,\ \ell = 1\}$ is called "2p", an so on.

It is important to notice that the lowest energy level possible is $n = 1$. This implies that the electron is not able to move to lower energy states. In this way, the state 1s (also called fundamental state) is stable and explains the presence of hydrogen in the Universe. Moreover, the 1s state exhibits spherical symmetry since the wave function is
$$\Psi_{n=0,\ell=0,m_\ell=0} = \frac{1}{\sqrt{\pi r_0^3}} e^{-\frac{r}{r_0}}, \tag{1.13}$$

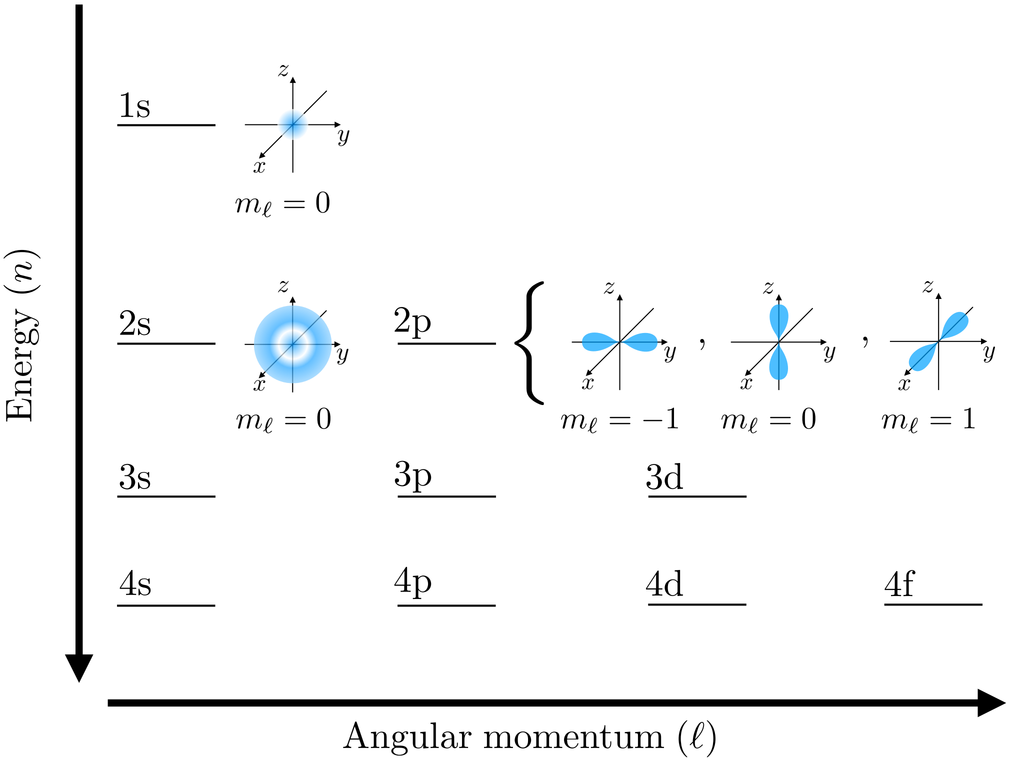


Fig. 1.3 Sketch of the quantum energy levels of the hydrogen atom. The orbitals for the two lowest energy levels are shown.

where r_0 is a scale parameter that determines the extension of the region where it is most likely to find the electron (orbital). In this state the wave function of the electron decreases monotonically when distancing from the proton.

The complexity of the wave functions augments with the quantum numbers n and ℓ (see Fig.1.3). For example, the orbital 2s still exhibits spherical symmetry, but the wave function does not depend monotonically on the distance to the proton anymore. Instead, the orbital has a maximum in $r = 0$, then the wave function falls, rises up to a hill and decreases monotonically until $r = \infty$. Additionally, the family of orbitals 2p ($\ell = 1$, $m_\ell = -1, 0, 1$) no longer presents spherical symmetry and the orbitals are divided in two elongated lobes along a direction determined by the value of m_ℓ.

Roughly speaking, n determine size of the orbital. In states with higher n the electron is less binded and the orbitals are wider. Meanwhile, the quantum number ℓ is a proxy for the eccentricity of the orbital. States with higher ℓ are more elongated along a given direction set by the magnetic quantum number m_ℓ.

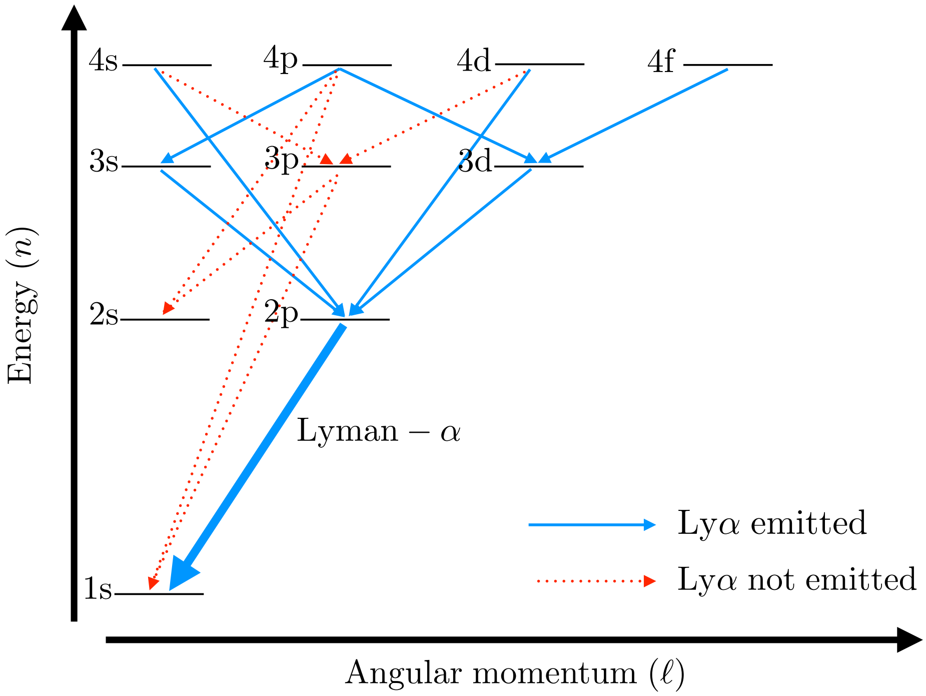


Fig. 1.4 Possible cascades in the hydrogen atoms after recombination assuming the dipolar approximation. In dotted red arrows we show the cascades that does not produce Lyα photons. Meanwhile, blue arrows indicate cascades in which the last transition is from the level $2p$ to $1s$, thus Lyα is produced. Figure taken from Dijkstra (2017).

1.2.3 Transition between quantum states

In quantum mechanics a bounded electron can change its state by absorbing or emitting photons to gain or lose energy, respectively. Additionally, the only stable state in the hydrogen atom is the fundamental. Thus, after a short time, an electron at the state $\{n, \ell\}$ transits to another state of lower energy $\{n', \ell'\}$ emitting a photon with energy $E_\gamma = E_{n'} - E_n$. The photons generated when electrons fall from the first excited energy level ($n = 2$) to the fundamental ($n = 1$) are called Lyman-α (from now on Lyα) in honor of their discoverer, Theodore Lyman in 1906. The corresponding wavelength to this transition is ~ 1215.67Å, which belongs to the ultraviolet part of the electro-magnetic spectrum.

In general, not all the possible transitions between different states are permitted. The *selection rules* of the allowed transitions are based on the conservation of linear momentum, energy and angular momentum. These depend on the physical regime considered. For example, in the dipole approximation it is assumed that the typical wavelength of the photon emitted during the transition is much larger than the size of the atom. In this case, there are two restrictions

1. The parity, $P = (-1)^\ell$ (even when $P = 1$ and odd when $P = -1$), of the final state must be different of the parity of the initial state. This translates into $\Delta\ell = odd$.
2. The angular momentum of the electron might change up to one unit ($\hbar$). This implies $\Delta\ell = 0, \pm 1$.

Therefore, allowed transitions in the dipole approximation must fulfill $\Delta\ell = \pm 1$. The other transitions are called *forbidden.* The quantum states that can only decay to the fundamental level through forbidden transitions, such as 2s, are called meta-stables. In fact, the forbidden transitions occur, but with much lower probability. For example, if an electron in the state 2s emits 2 photons it is allowed to transits to the state 1s.

Since the fundamental level is the only stable state, every electron in any different state $\{n, \ell\}$ will migrate to a lower energy state $\{n', \ell'\}$ radiating photons. This process is repeated until the electron reaches the fundamental level. These events are called *cascades.* In Fig.1.4 we show the allowed cascades by the dipole approximation in the hydrogen atom. On one hand, most of the cascades end up in the fundamental state. From these, a fraction comes directly from the state 2p, thus producing Lyα. On the other hand, some of the cascades get stuck at the state 2s, which has a very small probability of decaying to the ground level.

1.2.4 Relativistic corrections to the hydrogen atom

In the quantum description of the hydrogen atom presented in §1.2.2 we introduced the discrete energy levels as a solution of the time independent Schrödinger equation for a non-relativistic particle. In this formalism the energy of a given state only depends on the principal quantum number n, thus all the orbitals are degenerated (have the same energy). However, when some corrections are included in the anlysis the energy levels split. These corrections are:

- The kinetic energy of electrons given by special relativity

$$E_K = m_e c^2 \left(1 - \frac{1}{\sqrt{1-\beta^2}}\right), \tag{1.14}$$

where m_e is the rest frame electron mass, c is the speed of light in vacuum and $\beta = v_e/c$ where v_e is the electron velocity. For small velocities ($v_e \ll c$) this equation reproduces the classical description for the kinetic energy $\frac{1}{2}m_e v_e^2$.

- Spin-orbit coupling caused by the electromagnetic interaction between the Coulomb field generated by the proton and the magnetic dipole of the electron generated by its intrinsic angular momentum.
- The Darwin term that takes into account the rapid quantum fluctuations that the electron undergoes within the electric potential (Zitterbewegung effect). Due to the quantum oscillations of the electron around its mean position, the Coulomb potential is smoothed around the origin. This only affects to the $\ell = 0$ states and corrects that their wave functions vanish at $r = 0$.

When all these effect are combined the energy correction to the non-relativistic hydrogen atom is

$$\Delta E_{n,\ell,j} = \frac{E_n \alpha^2}{n}\left(\frac{1}{j+1/2} - \frac{3}{4n}\right), \tag{1.15}$$

where α is the fine structure constant and $j = |\ell \pm 1/2|$ is the total angular momentum quantum number. From now on, we will use the notation '$n\ell j$' for referring to fine structure levels. In this way, $2S_{1/2}$ denotes the quantum state $\{n,\ell,j\} = 2,0,1/2$. Then, each $\{n,\ell,j\}$ state is divided in $2j+1$ degenerated quantum states. Theses are referred with the projection of the total angular momentum quantum number

$$m_j = -j\ ,\ -j+\frac{1}{2}\ ,\ \dots\ ,\ j-\frac{1}{2}\ ,\ j \tag{1.16}$$

In this way, the states $P_{1/2}$ are split in two degenerated states, while the states $P_{3/2}$ into four.

The corrections presented in this section change the shape of the wave functions of the electrons with respect to the non-relativistic case. It will be of most interest later on that the shape of the $2p$ orbitals change (as illustrated in Fig. 1.5). In particular, the probability density P_θ is (White, 1934)

$$2P_{\frac{1}{2}}\left\{\ m_j = \pm\tfrac{1}{2} \quad\rightarrow\quad P_\theta = 1\right. \quad \text{and} \quad 2P_{\frac{3}{2}}\left\{\begin{array}{lll} m_j = \pm\frac{3}{2} & \rightarrow & P_\theta = \frac{3}{2}\sin^2\theta \\ m_j = \pm\frac{1}{2} & \rightarrow & P_\theta = \frac{1}{2}(3\cos^2\theta + 1) \end{array}\right. , \tag{1.17}$$

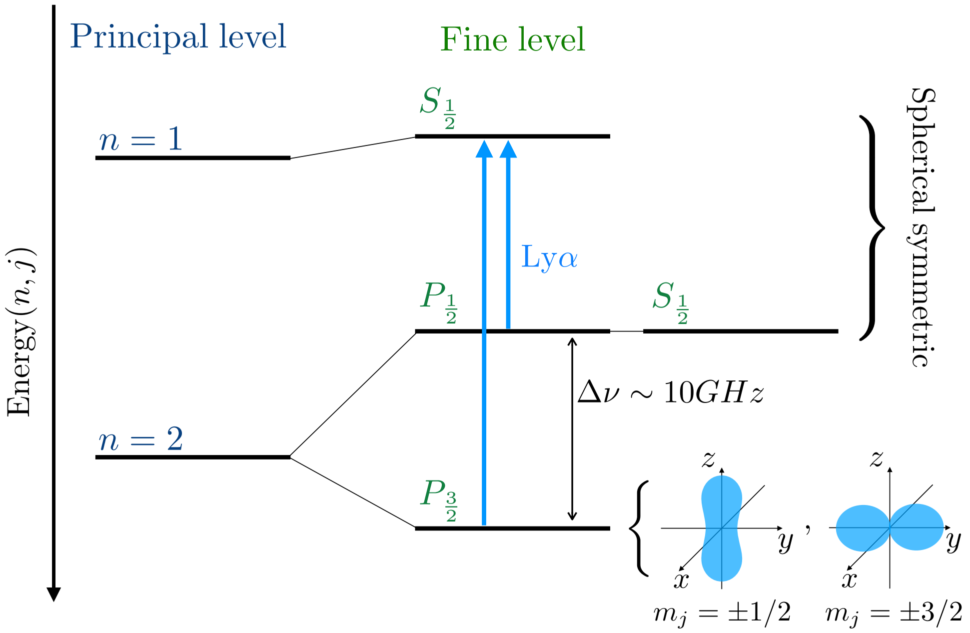


Fig. 1.5 Energy levels of the hydrogen atom fine structure up to $n = 2$. The orbitals of the states with $j = 1/2$ are spherical. Meanwhile, the orbitals of $j = 3/2$ exhibits revolution symmetry.

where θ is the subtended angle between a point and the projection axis of m_j. The states $2P$ are axially symmetric. However, the states $2P_{1/2}$ are spherically symmetric while the $2P_{3/2}$ orbitals exhibit the characteristic double lobe shape. This is shown in Fig.1.5.

1.3 Lyman-α physics

1.3.1 Lyman-α emission by a hydrogen cloud

Lyα photons are emitted when electrons in the $n = 2$ energy level decay to the fundamental level. However, electrons in higher energy levels might also produce Lyα photons through the cascades. In particular, the cascades in which the final transition is 2p→1s contribute to the Lyα production. There are two ways in which an electron can populate an excited energy level ($n > 1$) in a hydrogen cloud containing also free electrons.

First, the interactions between free electrons and hydrogen atoms. In this case a free electron transfer part of its kinetic energy to an electron bounded to a proton. In this way, the binded electron moves to a higher energy level and produces a cascade of photons. In practice, the kinetic energy of the electrons is transformed in radiation, which lowers the temperature of the gas. This physical process is also known as '*radiative cooling*' and for hydrogen it is most efficient at temperatures close to 10^4K. Radiative cooling plays an important role in the early stages of galaxy formation, as it helps to lower the gas temperature from 10^5K to 10^4K.

The second and most interesting for the topics covered in this thesis is *recombination*. Recombination is the physical process by which a proton captures a free electron with kinetic energy E_k and a photon with energy $E_\gamma = E_K - E_n$ is emitted, where E_n is the energy of the state in which the electron is binded. This constitutes the main mechanism by which a star forming galaxy produces Lyα photons (as seen later).

In general, the probability of a recombination event resulting in Lyα emission depends on:

1. The probability that a recombined electron get binded in the state $\{n,\ell\}$. This is given by
$$P_{n,\ell}(T) = \frac{\alpha_{n,\ell}(T)}{\sum\limits_{n'=n_{\rm min}}^{\infty} \sum\limits_{\ell'=0}^{n'-1} \alpha_{n',\ell'}(T)} = \frac{\alpha_{n,\ell}(T)}{\alpha_{\rm tot}(T)}, \tag{1.18}$$
where $\alpha_{n,\ell}(T)$ is the recombination coefficient of the quantum state $\{n,\ell\}$ defined as the number of recombinations in the state $\{n,\ell\}$ per unit of time, volume, number density of free electrons and protons. Meanwhile, $\alpha_{\rm tot}(T)$ is the total recombination coefficient, which measures the specific number of recombinations per volume and time. A full calculation of these can be found in Burgess (1965).

2. The probability $P(\{n,\ell\} \to ... \to 2p \to 1s)$ that a cascade starting at the state $\{n,\ell\}$ ends in a $2p \to 1s$ transition. This is given by
$$P(\{n,\ell\} \to \text{Ly}\alpha) = \sum_{n',\ell'} P(\{n,\ell\} \to \{n',\ell'\}) \times P(\{n',\ell'\} \to \text{Ly}\alpha), \tag{1.19}$$
where $P(\{n,\ell\} \to \{n',\ell'\})$ is the probability of the state $\{n,\ell\}$ decaying in $\{n',\ell'\}$, where $\{n',\ell'\}$ is a intermediate states between $\{n,\ell\}$ and 2p. This is defined as
$$P(\{n,\ell\} \to \{n',\ell'\}) = \frac{A_{\{n,\ell\},\{n',\ell'\}}}{\sum\limits_{n'',\ell''} A_{\{n',\ell'\},\{n'',\ell''\}}}, \tag{1.20}$$

where $A_{\{n,\ell\},\{n',\ell'\}}$ is the Einstein coefficient for spontaneous emission in the transition $\{n,\ell\} \to \{n',\ell'\}$, which is defined as the probability per unit of time of that decay to happen. The $A_{\{n,\ell\},\{n',\ell'\}}$ values depend on the regime considered. For example, in the dipole approximation $P(2s \to 1s) = 0$ and $P(2p \to 1s) = 1$ since only $\Delta\ell = \pm 1$ transitions are allowed.

All together, the probability that a recombination event produce a Lyα photon is

$$P(\mathrm{Ly}\alpha) = \sum_{n=n_{\mathrm{min}}}^{\infty} \sum_{\ell=0}^{n-1} P_{n,\ell}(T) \times P(\{n,\ell\} \to \mathrm{Ly}\alpha). \tag{1.21}$$

The Lyα rate production by recombination depends on the physical properties of the neutral hydrogen cloud through the recombination coefficients $\alpha_{n,\ell}(T)$. In particular, usually there are two different scenarios contemplated:

- Case-A recombination : The medium is optically thin to all frequencies. In this case the lowest energy level where a photon can be directly binded is the fundamental ($n_{\mathrm{min}} = 1$).
- Case-B recombination : The medium is optically thick to ionizing photons, i.e., photons able to dissociate the hydrogen atom when the electron is at any quantum state. Thus, ionizing photons have energy $E_\gamma \geq |E_{n=1}|$. In general, an ionizing photon is emitted every time a proton captures an electron to the ground state, since $E_\gamma = E_k - E_{n=1}$ is strictly greater than $|E_{n=1}|$. Then, this ionizing photon photoionize another hydrogen atom, creating a new pair of free proton and electron, where the new electron kinetic energy E'_k is the same as the kinetic energy of the photon that was recombined E_k. In case-B recombination this process take a very short amount of time since the medium is optically thick to this radiation. Therefore, in this frame, all the recombination events to the ground level are considered cancelled out, i.e., α_{1s} is null and $n_{\mathrm{min}} = 2$.

The probability of emitting a Lyα photon after a recombination event is approximately

$$P_A(\mathrm{Ly}\alpha) = 0.41 - 0.165 \log\left(\frac{T}{10^4}\right) - 0.015 \left(\frac{T}{10^4}\right)^{-0.44} \tag{1.22}$$

in the case-A recombination (Dijkstra, 2014) and

$$P_B(\mathrm{Ly}\alpha) = 0.686 - 0.106 \log\left(\frac{T}{10^4}\right) - 0.009 \left(\frac{T}{10^4}\right)^{-0.44} \tag{1.23}$$

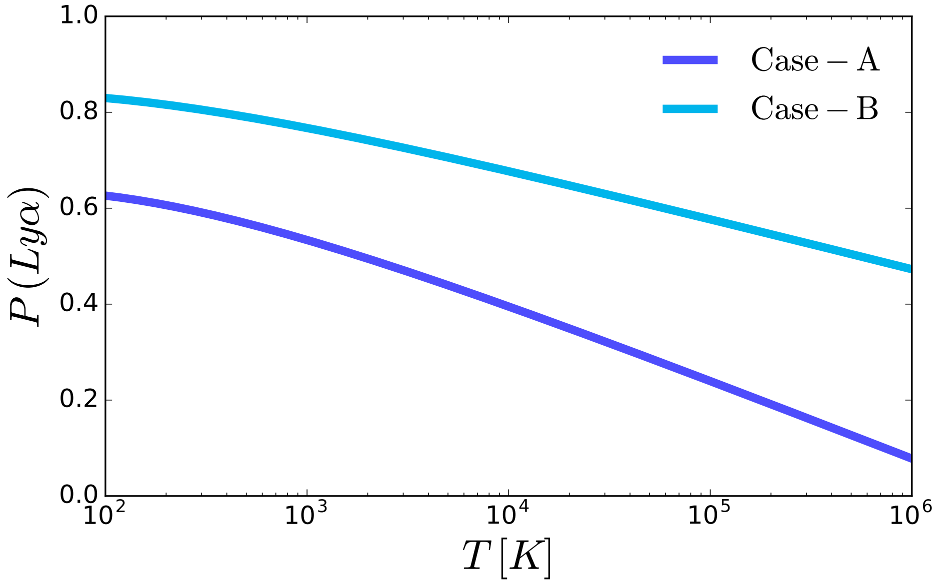


Fig. 1.6 Probability of emitting a Lyα photon after recombination in case-A (purple) and case-B (blue) as a function of the gas temperature. Figure taken from Dijkstra (2017).

in case-B recombination (Cantalupo et al., 2008). In Fig.1.6 it is displayed the probability of emitting a Lyα photon after a recombination. At a given temperature, the probability of emitting Lyα photons after a recombination is higher in case-B, since the total recombination specific ratio $\alpha_{\rm tot}$ is lower. Additionally, case-A varies more quickly with temperature compared to case B.

1.3.2 Lyman-α - hydrogen cross-section

Lyα photons have a high chance to interact with neutral hydrogen atoms. In particular, the cross-section close to the Lyα wavelength $\lambda_{\rm Ly\alpha}$ between a Lyα photon and a single $\rm H_I$ atom for a given frequency ν (Peebles, 1993) is

$$\sigma_{\rm Ly\alpha\times H_I}(\nu) = \frac{3\lambda^2_{\rm Ly\alpha}}{8\pi}\frac{A^2_{\rm Ly\alpha}(\nu/\nu_{\rm Ly\alpha})^4}{(2\pi)^{-2}(\nu-\nu_{\rm Ly\alpha})^2+(A_{\rm Ly\alpha}/2)^2(\nu/\nu_{\rm Ly\alpha})^6} \tag{1.24}$$

where $\nu_{\rm Ly\alpha}$ is the Lyα frequency and $A_{\rm Ly\alpha} \equiv A_{\{2p\},\{1s\}} = 6.25\ 10^8 s^{-1}$ is the Einstein coefficient for the Lyα transition. This cross-section (shown in Fig.1.7) exhibits a prominent peak at the Lyα wavelength, where it reaches a value as high as $\sim 3\ 10^{-12} cm^{-2}$. When the wavelength is shifted a little bit from Lyα, $\sigma_{\rm Ly\alpha\times H_I}$ decreases 7 orders of mag-

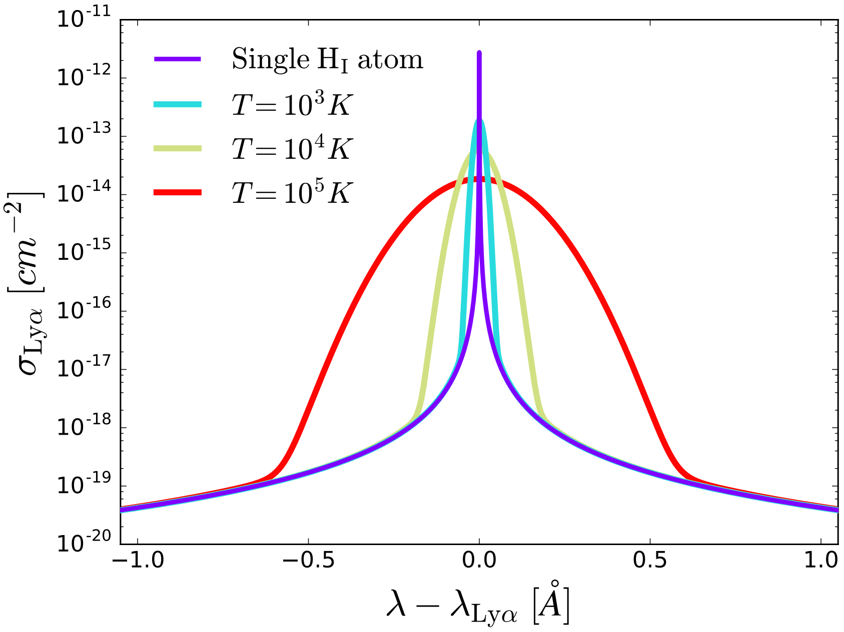


Fig. 1.7 Cross-section between a Lyα photon and i) single $\mathrm{H_I}$ atom (purple) ii) an $\mathrm{H_I}$ could at $T = 10^3 K$ (blue), $T = 10^4 K$ (yellow) and $T = 10^5 K$ (red).

nitude in roughly 0.25Å. Then it progressively keeps getting smaller, but at a slower rate.

In astrophysical environments, hydrogen atoms are not alone, instead they form big $\mathrm{H_I}$ clouds. Therefore, it is also interesting to study the cross-section of Lyα photons and macroscopic $\mathrm{H_I}$ gas, $\sigma_{\mathrm{Ly}\alpha}$. In fact, $\sigma_{\mathrm{Ly}\alpha}$ depends on the $\mathrm{H_I}$ region temperature due to three main reasons :

- The cross-section of a single hydrogen atom and Lyα photons depends on the frequency.
- Particles constituting a gas at a given temperature T have kinetic energy. If we assume an ideal gas behaviour, then the velocity modulus v of the individual atoms distributes as a Maxwell–Boltzmann distribution, i.e,

$$F_{\mathrm{MB}}(v) = 4\pi \left(\frac{m_H}{2\pi k_B T} \right)^{3/2} v^2 e^{-\frac{m_H v^2}{2 k_B T}}, \tag{1.25}$$

where k_B is the Boltzmann constant and m_H is the rest frame mass of the hydrogen atom. In this approximation, the orthogonal components of the velocity, v_i, are independent and distributed as

$$f_{\mathrm{MB}}(v_i) = \left(\frac{m_H}{2\pi k_B T}\right)^{1/2} e^{-\frac{m_H v_i^2}{2k_B T}}. \tag{1.26}$$

- Due to the relativistic Doppler effect, in the rest frame of a particle, photons emitted at a given frequency ν are received at the frequency ν', such as

  $$\frac{\nu'}{\nu} = 1 - \frac{v_{\parallel}}{c} \tag{1.27}$$

  where $v_{\parallel}$ is the relative velocity between the particle and the source of photons along the direction joining them (line of sight) and c is the speed of light. Therefore, photons emitted by a source approaching the particle are received at a higher frequency (blueshifted) than the frequency at which they were originally emitted. Meanwhile, if the source and the particle are distancing the photons are received at a lower frequency (redshifted).

The combination of these effects causes that individual $\mathrm{H_I}$ atoms in a cloud have different probabilities of interacting with Lyα depending on their own velocity when they receive the photon. Thus, the cross-section between Lyα photons and $\mathrm{H_I}$ clouds is the convolution of the 3D velocity $\vec{v}$ distribution of $\mathrm{H_I}$ atoms and the cross-section of single $\mathrm{H_I}$ atoms with Lyα, i.e,

$$\sigma_{\mathrm{Ly}\alpha}(\nu, T) = \int \sigma_{\mathrm{Ly}\alpha\times\mathrm{H_I}}(\nu|\vec{v}) F_{\mathrm{MB}}(\vec{v})\ d^3\vec{v} \tag{1.28}$$

where $\sigma_{\mathrm{Ly}\alpha\times\mathrm{H_I}}(\nu|\vec{v})$ is the cross-section between single $\mathrm{H_I}$ atoms with velocity $\vec{v}$ and Lyα photons at a given frequency. It is useful to decompose $\vec{v}$ in orthogonal components parallel ($v_{\parallel}$) and perpendicular ($v_{\perp,1}$, $v_{\perp,2}$) to the line of sight. Thus,

$$\sigma_{\mathrm{Ly}\alpha}(\nu, T) = \int_{-\infty}^{\infty} f_{\mathrm{MB}}(v_{\perp,1})\ dv_{\perp,1} \int_{-\infty}^{\infty} f_{\mathrm{MB}}(v_{\perp,2})\ dv_{\perp,2} \int_{-\infty}^{\infty} \sigma_{\mathrm{Ly}\alpha\times\mathrm{H_I}}(\nu|v_{\parallel}) f_{\mathrm{MB}}(v_{\parallel})\ dv_{\parallel} \tag{1.29}$$

where the integrals of $v_{\perp,1}$ and $v_{\perp,2}$ are 1 since $\sigma_{\mathrm{Ly}\alpha\times\mathrm{H_I}}$ only depends on the motion along the line of sight (Doppler effect). Now, if we substitute Eq.1.24, 1.26 and 1.27, then $\sigma_{\mathrm{Ly}\alpha}$ can be written as (Orsi et al., 2012)

$$\sigma_{\mathrm{Ly}\alpha}(x) = \frac{3\lambda_{\mathrm{Ly}\alpha} a_V^2}{2\pi^{3/2}} \int_{-\infty}^{\infty} \frac{e^{-y^2}}{(y-x)^2 + a_V} dy, \tag{1.30}$$

where we have defined the Voight parameter a_V as

$$a_V = \frac{A_{\mathrm{Ly}\alpha}/4\pi}{\Delta\nu_D} = 4.7 \times 10^{-4} \left(\frac{T}{10^4}\right)^{-1/2}, \tag{1.31}$$

with $v_{th} = \left(\frac{2k_B T}{m_H}\right)^{1/2}$ and $\Delta\nu_D = \nu_{\mathrm{Ly}\alpha}\frac{v_{th}}{c}$. Additionally, x is the frequency in Doppler units, defined as

$$x \equiv \frac{\nu - \nu_{\mathrm{Ly}\alpha}}{\Delta\nu_D}. \tag{1.32}$$

It is common to rewrite the cross-section of Lyα and a $\mathrm{H_I}$ cloud as

$$\sigma_{\mathrm{Ly}\alpha}(x,T) = 5.686 \times 10^{-14} \left(\frac{T}{10^4}\right)^{-1/2} \frac{H(a_V,x)}{\sqrt{\pi}}, \tag{1.33}$$

where

$$H(a_V,x) = \frac{a_V}{\pi}\int_{-\infty}^{\infty} \frac{e^{-y^2}}{(y-x)^2 + a_V} dy = \begin{cases} \sim e^{-x^2} & \text{core} \\ \sim \frac{a_V}{\sqrt{\pi} x^2} & \text{wings} \end{cases}, \tag{1.34}$$

is the Voight profile (Hjerting, 1938). In fact, computing this integral can be computationally expensive, specially if it has to be evaluated several times. Tepper-García (2006) calculated an approximation for the Voight profile :

$$H(a_V,x) = e^{-x^2} - \frac{a_V}{\sqrt{\pi}x^2}\left[e^{-2x^2}\left(4x^4 + 7x^2 + 4 + \frac{3}{2x^2}\right) - \frac{3}{2x^2} - 1\right]. \tag{1.35}$$

In Fig.1.7 we compare the cross-section between Lyα and a single $\mathrm{H_I}$ atom and with $\mathrm{H_I}$ clouds at different temperatures. In general, $\sigma_{\mathrm{Ly}\alpha \times \mathrm{H_I}}$ at exactly the Lyα wavelength is higher than $\sigma_{\mathrm{Ly}\alpha}$ at all temperatures. Then, the cross-section with $\mathrm{H_I}$ clouds exhibits a bell shape which size increases with temperature and finally it matches the cross-section of a single $\mathrm{H_I}$ atom. A simple interpretation is that, since $\mathrm{H_I}$ atoms present random peculiar motions, the number of $\mathrm{H_I}$ atoms that would received Lyα photons exactly at Lyα frequency ($v_\parallel \sim 0$) is reduced, which effectively reduces $\sigma_{\mathrm{Ly}\alpha}$ close to Lyα. However, this augments $\sigma_{\mathrm{Ly}\alpha}$ in the range where it is probable to find a $\mathrm{H_I}$ atom with velocity ($|v_\parallel| \gg 0$) such as a photon close to Lyα is perceived as Lyα. This is causing the bell behaviour previously discussed. Additionally, as the temperature increases, this range becomes broader and the bell is extended. Finally, when there are no chances of finding a $\mathrm{H_I}$ atom that would received a photon as if it was at Lyα, both cross-sections behave similar ($\sigma_{\mathrm{Ly}\alpha\times\mathrm{H_I}} \sim \sigma_{\mathrm{Ly}\alpha}$).

1.3.3 Lyman-α transmission

The optical depth of Lyα photons through a $\mathrm{H_I}$ cloud at a given temperature is defined as

$$\tau(\nu) = \int_0^\infty n_{\mathrm{H_I}}(s)\sigma_{\mathrm{Ly}\alpha}(\nu, V_{\mathrm{H_I}}(s))\; ds \tag{1.36}$$

where $n_{\mathrm{H_I}}(s)$ is the $\mathrm{H_I}$ number density at given distance s from the source. Additionally, the term $\sigma_{\mathrm{Ly}\alpha}(\nu, V_{\mathrm{H_I}}(s))$ contemplates the possibility that the $\mathrm{H_I}$ cloud might exhibit a bulk velocity that may also depend on s. If we consider that the $\mathrm{H_I}$ cloud is at rest with respect to the Lyα source,

$$\tau(\nu) = \mathrm{N_H} \;\times\; \sigma_{\mathrm{Ly}\alpha}(\nu) \tag{1.37}$$

where $\mathrm{N_H}$ denotes the $\mathrm{H_I}$ column density.

Hydrogen clouds scatter Lyα photons very efficiently. To illustrate this, let us define the transmission as the fraction of Lyα photons that do no interact with the $\mathrm{H_I}$ cloud, i.e.,

$$T(\nu) = e^{-\tau(\nu)}. \tag{1.38}$$

In Fig.1.8 we show the transmission of Lyα through $\mathrm{H_I}$ regions with different column densities for a fixed temperature $T = 10^4 K$. The transmission is minimum at Lyα and it decreases symmetrically very fast with $\mathrm{N_H}$. In fact, for $\mathrm{N_H} = 10^{12} cm^{-2}$ only a very small fraction of photons closer than 0.1Å to Lyα would interact. Meanwhile, basically, all photons closer than 0.1Å to Lyα are scattered by the cloud if $\mathrm{N_H} = 10^{16} cm^{-2}$. The $\mathrm{H_I}$ column density expected inside galaxies is about $10^{20} cm^{-2}$ and $10^{22} cm^{-2}$. For these values of $\mathrm{N_H}$, all the photons closer than 1Å and 10Å respectively are expected to be scattered. The wings of absorption go even further and photons at $\lambda - \lambda_{\mathrm{Ly}\alpha} < 10$Å and < 100Å respectively are likely to interact with the cloud.

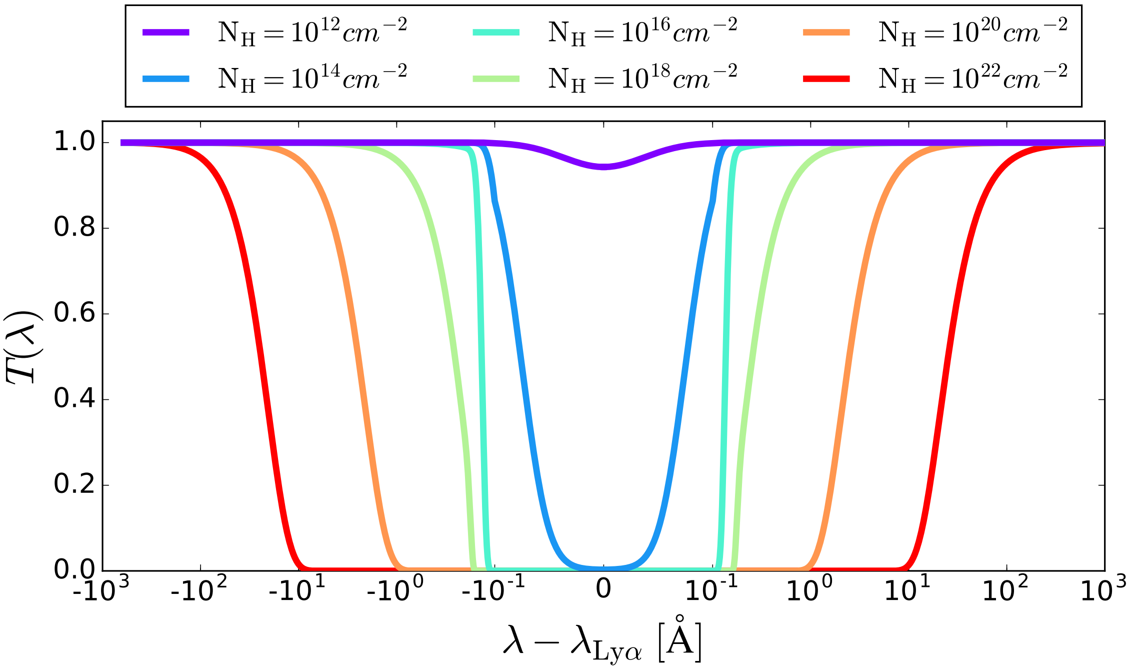


Fig. 1.8 Transmission around Lyα for a H_I gas at $T = 10^4 K$ and several column densities.

1.3.4 Lyman-α interaction

In this section we explain the two most common Lyα interactions in a astrophysical environment: with H_I and dust grains.

Interaction with hydrogen

When a H_I atom with an electron in the ground level absorbs a Lyα photon its electron is excited to the $2p$ orbital. Then, after a short time, the electron decays to the fundamental level by emitting another Lyα photon. As shown in §1.2.4 the $2p$ orbital is divided in two $2P_{1/2}$ and four $2P_{3/2}$ states. The processes $1S_{1/2} \rightarrow 2P_{1/2} \rightarrow 1S_{1/2}$ and $1S_{1/2} \rightarrow 2P_{3/2} \rightarrow 1S_{1/2}$ are usually referred as K-line and H-line, respectively. These spectral lines are separated by less than 0.01Å.

The direction of the new Lyα photon is determined by the phase function, which depends on the quantum structure of the hydrogen atom. The quantum state $2P_{1/2}$ exhibits spherical symmetry while the states $2P_{3/2}$ are only axially symmetric. This causes that the phase function of the K-line is spherically symmetric. Meanwhile,

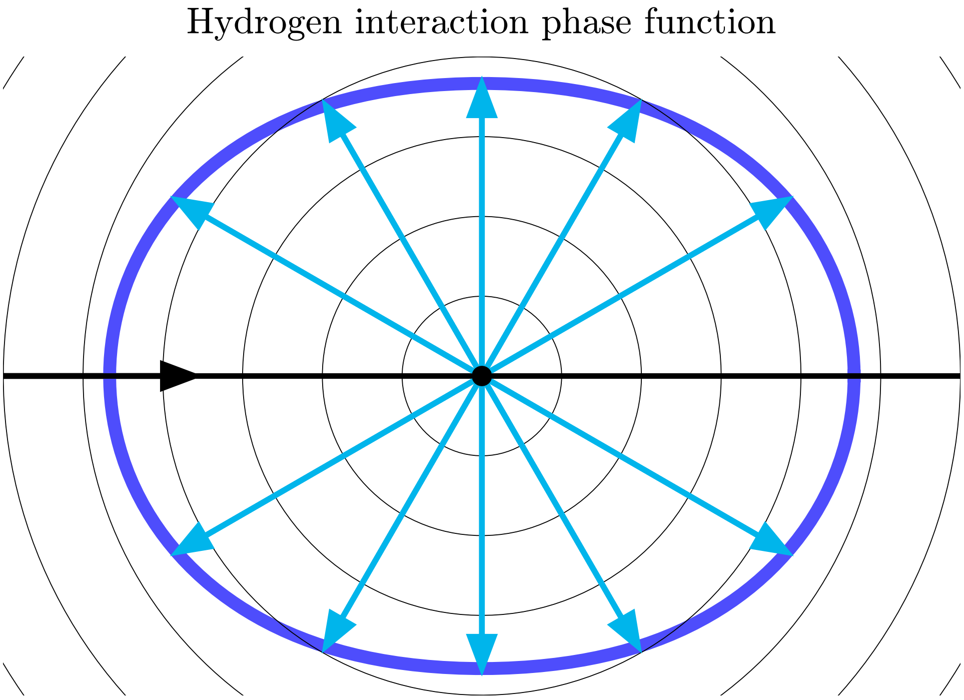


Fig. 1.9 Phase function of the interaction between a Lyα photon and a H_I atom (purple). The black horizontal line indicates the incident direction. The black dot indicates the position of the H_I atom. The thin circular lines are just circumferences around the H_I atom. The blue arrows indicate the emissivity in that direction. The larger the arrow are, the more photon are sent in that direction.

photons in the H-line are polarized and their phase function is a dipole. Therefore, the total phase function of a Lyα-H_I interaction is a superposition of isotropic and dipolar scatter. The particular weight of each component depends on the probability of producing Lyα photons through the H-line or the K-line. In fact, the total phase function of a Lyα-H_I interaction as at the fine structure level is (Brasken and Kyrola, 1998)

$$p(\theta) = \frac{11}{12} + \frac{3}{12}\cos^2\theta, \tag{1.39}$$

where θ is the angle between the incident and the emitted photon. In Fig.1.9 we show the phase function of an interaction with hydrogen $p(\theta)$. The probability of sending the Lyα photon forwards is the same than to scattering it backwards. Additionally, the probability that the new Lyα photon is emitted perpendicular to the incident direction is smaller (but not null) than to emit the photon closer to the line of incident direction.

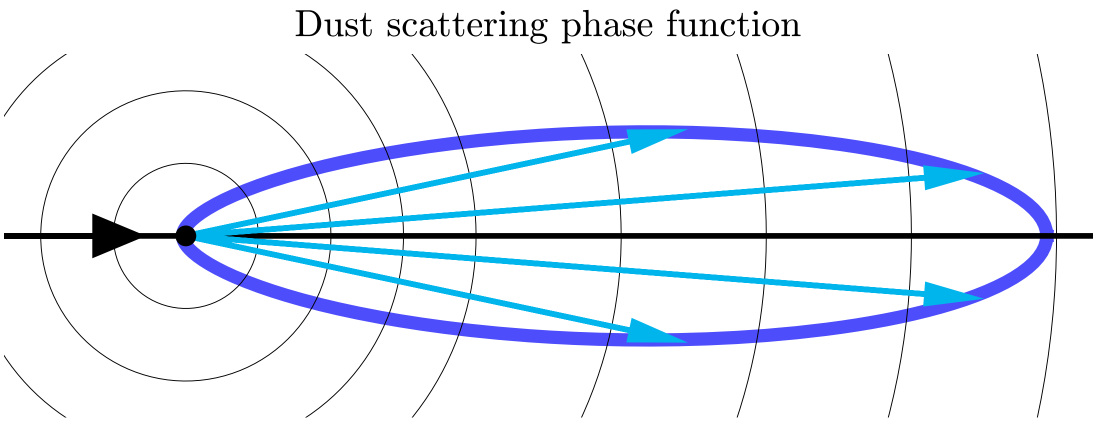


Fig. 1.10 Same as Fig. 1.9 but for the interaction between a Lyα photon and a dust grain.

In general, the Lyα source and the hydrogen atom will have a relative velocity. This causes that the Lyα is received Doppler shifted by the H_I atom. This affects to the chances of interacting, as the cross-section depends on frequency. Then, the new photon in emitted at Lyα in the H_I rest frame. This is called coherent scattering. The new photon is perceived by the initial Lyα source Doppler shifted from Lyα. In particular, if the H_I is getting away from the initial Lyα source, then the radiation is received redshifted. Meanwhile, if the H_I atom was approaching the initial source, it is blueshifted.

Interaction with dust

When a Lyα photon interacts with a dust grain there are two possible outcomes. The photon is either i) absorbed or ii) scattered. This depends on the albedo of the dust. Typical value of the dust albedo at the Lyα wavelength is 0.4 (Orsi et al., 2012).

In general, if the photon is not absorbed its direction change. The typical phase function for photons (not only Lyα) interacting with dust is (Henyey and Greenstein, 1941)

$$p(\theta) = \frac{\gamma(1-g^2)}{4\pi} \frac{1}{(1+g^2-2g\cos\theta)^{3/2}}, \tag{1.40}$$

where γ is the spherical albedo and g is the asymmetry parameter that determines the shape of $p(\theta)$. The spherical symmetry is recovered when $g = 0$, as the dependence with $\cos\theta$ vanish. Then, in the extreme case where $g = -1$ ($g = +1$) all the radiation is sent backwards (forwards). At Lyα, $g \sim 0.73$. In Fig.1.10 we show the phase function for a Lyα photon and dust. $p(\theta)$ is very elongated along the incident direction since g takes a great value at Lyα. Additionally, the probability that the photon gets scattered backwards or perpendicular to the incident direction is very small. Indeed, in most cases the photon will be scattered forward and very close to the incident direction.

1.4 Lyman-α emitters

Lyman-α emitters are extragalactic sources exhibiting strong Lyα emission in their spectrum. This population is mainly constituted by *Quasi Stellar Objects* (QSOs) and galaxies with ongoing star formation. On one hand, in QSOs the Lyα photons are generated by the interaction of black holes and their accretion disks. On the other hand, in star forming galaxies these photons come from the H_{II} regions hosting the newborn stars. This thesis is focused on the latter kind of Lyman-α emitter, that, from now on, will be referred to as LAE.

In this section we make an historical and theoretical introduction to LAEs, as well as to their perks and disadvantages.

1.4.1 First predictions

Partridge and Peebles (1967) suggested for the time that galaxies in an early state could emit Lyα photons.

> *"It seems possible that the Lyman-α line might be detected if it is a strong feature of the spectra of young galaxies."*
>
> Partridge and Peebles (1967)

The authors also suggested that the study of these galaxies could shed light on the cosmology of the Universe and the galaxy formation and evolution processes, opening a new window to the high redshift Universe.

The agents responsible for the Lyα photon production inside galaxies are the star forming regions (illustrated in Fig.1.11). In these regions the hot intergalactic gas (H_I at $\sim 10^4 K$) is cooled and transformed into molecular hydrogen clouds, which collapses

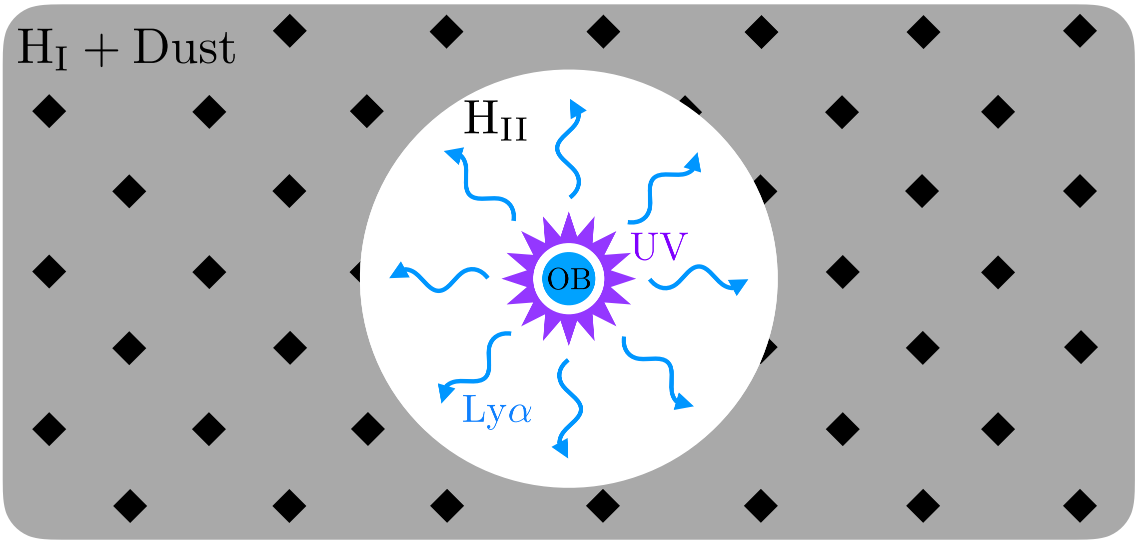


Fig. 1.11 Artistic representation of a star forming region where Lyα is generated. Young OB-*type* stars emit hard UV photons (purple) and ionize their surrounding cold H_I (grey) creating a ionized bubble (white). In the bubble, recombination episodes take place and Lyα (blue arrows) are emitted.

and form stars. The new stars generate radiation that heats the surrounding H_2 that gets dissociated. At first approximation, the spectral energy distribution (SED) of the radiation emitted by stars is as a black body, whose characteristic temperature is proportional to the mass of the star. This causes that very massive stars (*O-type* and *B-type*) emit a significant fraction of their photons in the ultraviolet. Very high energy photons ($\lambda < 912$Å) ionize the surroundings of the stars, forming H_{II} regions. In this hot plasma electrons and protons are split and recombined in H_I constantly, producing Lyα photons through radiative cascades.

Some of the emitted ionizing photons escape the galaxy, others interact with the H_I. Each of these interactions produces a pair of electron and proton. For each of these pairs a radiative cascade takes place. Thus, the number of Lyα photons generated is

$$N_{Ly\alpha} = (1 - f_{\rm esc}^{\rm ion}) \times N_\gamma \times P(\mathrm{Ly}\alpha) \tag{1.41}$$

where $f_{\rm esc}^{\rm ion}$ is the escape fraction of ionizing photons, N_γ is the number of emitted ionizing photons and P(Lyα) is the probability of emitting a Lyα after a recombination (Eq. 1.22 and 1.23). In this way, the produced Lyα luminosity is directly related to the star formation rate of the galaxy. The greater the amount of gas is transformed into stars, the higher is N_γ and therefore the produced Lyα flux. Additionally, Lyα is the transition with the highest probability in a radiative cascade at typical galactic temperatures, which makes Lyα the strongest hydrogen emission line. The high luminosity of the Lyα line is what makes possible to detect LAEs at cosmological distances.

An additional perk of using LAEs as a cosmological tool is the wavelength of Lyα, ~ 1215.67Å. While the Earth's atmosphere is opaque to Lyα at low redshift, it becomes transparent from $z \sim 2$ to $z \sim 8$. This makes LAEs a good population to understand the early stages of galaxy formation and evolution and the development of the Universe at these epochs.

1.4.2 Lyman-α radiative transfer

Some years after their hypothesized detection the astrophysical community started to realised the complexity of the LAEs. Harrington (1973) made a theoretical sutdy about the behaviour of resonant lines (such as Lyα) in an optically thick medium. They found that, if the medium has any photon destruction mechanism, the escape fraction of photons is reduced. Additionally, they also found that the surface brightness profile is modified by radiative transfer processes.

We have illustrated in Fig.1.12 the complex radiative transfer processes that Lyα go through in a H_I cloud containing dust. In this picture, monochromatic radiation at Lyα frequency (bottom left subpanel) is penetrating a H_I region. After traveling freely some distance the high cross-section between Lyα and H_I atoms makes them interact. In this interaction, the atom captures a Lyα photon traveling in a given direction. Then, the electron is excited and rapidly decays to the ground level emitting a Lyα photon in a different direction. Since the probability of being emitted forwards and backwards is the same (Eq.1.39), the length of the trajectory of Lyα is highly increased. See for example photon "c" of Fig.1.12.

In general, there are two main processes than might cause a flux reduction after entering into a H_I cloud:

- Photons that initially were traveling to the observer change their trajectory by scattering events and never reach it. This is illustrated with photon "a".
- A dust grain absorbs the Lyα photon (photon "c"). As the length of the trajectory of Lyα in the medium is augmented, the probability of interacting with dust also

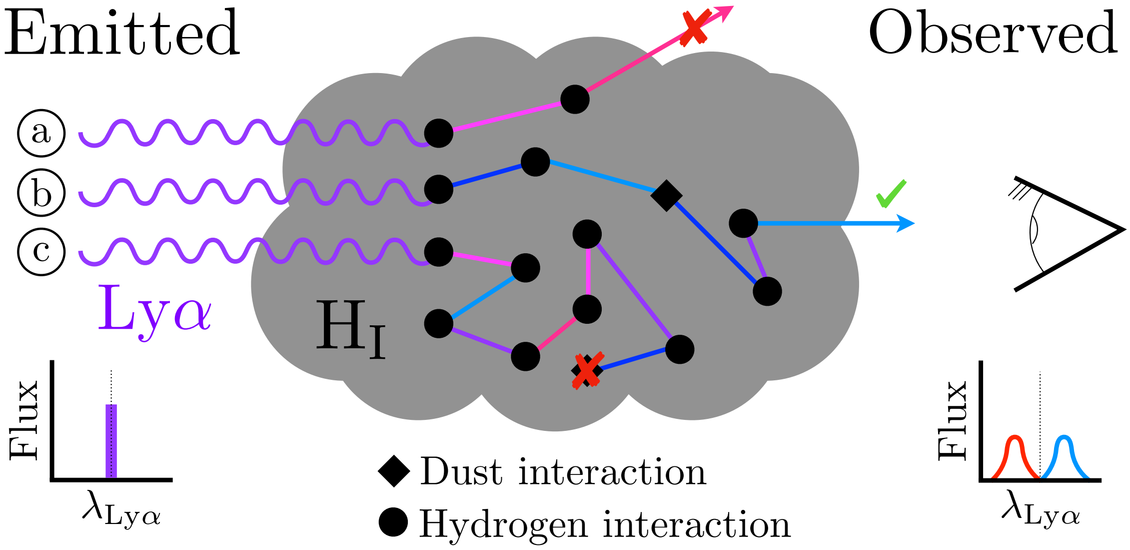


Fig. 1.12 Sketch of the Lyα radiative transfer processes that take place inside dusty H_I clouds. The frequency of the photons is color coded. First, all the photons enter into the cloud at Lyα. There, they interact with dust or H_I atoms. This changes their frequency and direction. Some radiation manage to escape and are observed with a different spectrum. Other photons are lost due to i) dust absorption or ii) they are scatter out of the line of sight.

increases. In fact, the scatter efficiency is so high that Lyα photons are extremely sensitive to dust.

The combination of these effect causes that the transmission of Lyα photons depends on the gas geometry, its optical depth and dust content. Moreover, since the final direction of these photons is modified, the surface brightness profile of the object is extended.

Additionally, in each scattering event the frequency of the Lyα photon changes slightly due to the peculiar motion of H_I atoms (see §1.3.4). This helps Lyα photons to escape, as consecutively scattering events Doppler shift more and more these photons until the medium becomes transparent to them. Therefore, the emerging Lyα line profile is also modified. Typically, the line profile exhibit a bimodal distribution. The blue part of the distribution comes from interactions with H_I atoms that, by their thermal peculiar motion, were moving towards the Lyα source at the moment of the scattering. In the same way, the red peak is caused by interactions with atoms that were distancing from the Lyα source.

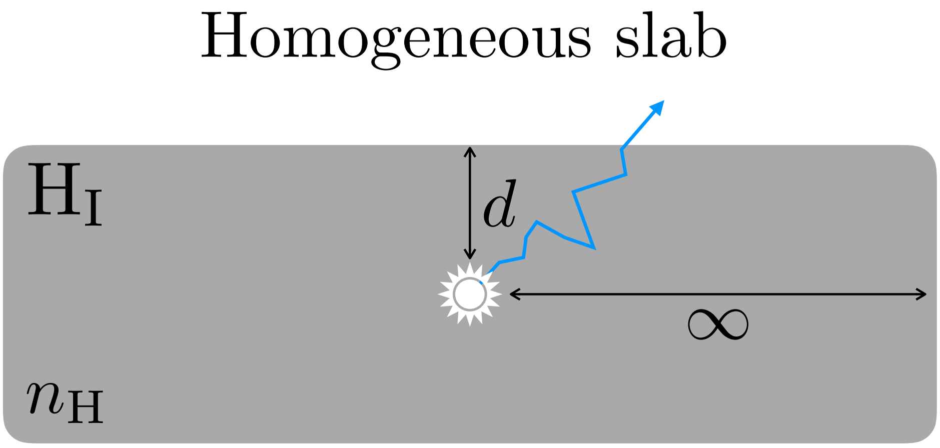


Fig. 1.13 Sketch of the homogeneous slab used in Neufeld (1990). In white we show the central Lyα source. One example of an escaping Lyα is shown in blue.

In Neufeld (1990) they provided analytic solutions to the complex radiative transfer of Lyα in an homogeneous static H_I slab at a certain temperature containing also dust. The slab is illustrated in Fig.1.13. This gas geometry is motivated by the astrophysical scenario in which Lyα photons are generated in narrow and elongated galaxy disks. In this way, the slab is considered infinite along its parallel direction and thin in the perpendicular. The Lyα source is located in the middle of the slab. Additionally, the H_I column density of this geometry is computed as

$$\mathrm{N_H} = n_\mathrm{H} \times d \tag{1.42}$$

where n_H is the number density of H_I atoms in the slab and d is half of the slab width. If monochromatic emission is assumed, then the fraction of Lyα photons that escape from the slab is

$$f_\mathrm{esc}^{\mathrm{Ly}\alpha} = \frac{1}{\cosh(\xi' \sqrt{(a_V \tau_0)^{1/3} \tau_a})} \tag{1.43}$$

where a_V is the Voight parameter (Eq.1.31), τ_0 is the optical depth in at Lyα(Eq.1.37), τ_a is the dust optical depth and $\xi' = \sqrt{3}/(\xi \pi^{5/12})$ where $\xi \sim 0.525$ is a free parameter.

In Fig.1.14 we show the escape fraction of Lyα photons in an homogeneous slab. We made this calculation for several values of H_I column density and metalliciy of the gas Z, where we assumed that the dust optical depth can be written as

$$\tau_a = (1 - A_{\mathrm{Ly}\alpha}) \frac{E_\odot}{Z_\odot} \mathrm{N_H} Z, \tag{1.44}$$

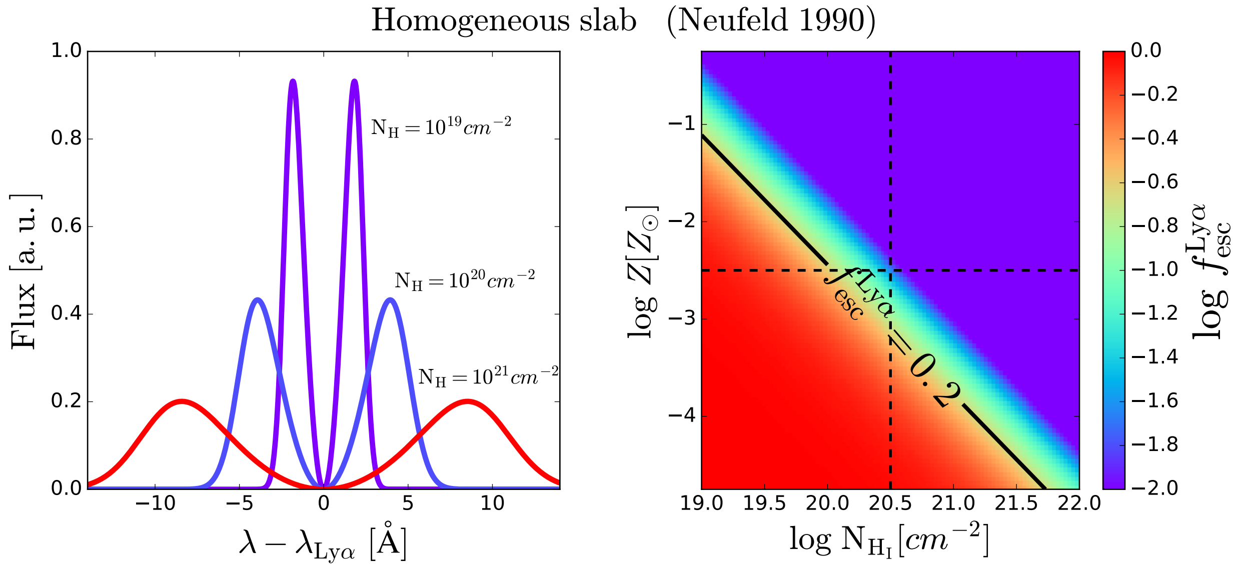


Fig. 1.14 Properties of the Lyα radiation after escaping from an homogeneous slab. Left: emerging line profiles as a function of the neutral hydrogen column density. Right: escape fraction as a function of metallicity and column density. The dark line shows $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha} = 0.2$

where $E_\odot = 1.77 \times 10^{-21} cm^{-2}$ is the ratio $\tau_a/\mathrm{N_H}$ for solar metallicity, $Z_\odot = 0.02$ (Granato et al., 2000) and $A_{Ly\alpha} = 0.39$ is the dust albedo at the Lyα frequency. In general, $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ decreases when the metallicity increases as the medium becomes more dusty. Moreover, when the $\mathrm{H_I}$ column density increases, the length of the trajectory of Lyα inside the medium is boosted and $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ decreases. Finally, for the typical values of metallicity ($Z \sim 10^{-2.5} Z_\odot$) and column density ($\mathrm{N_H} \sim 10^{20.5} cm^{-2}$) the escape fraction of Lyα photon is close to 10^{-2}. This means that the Lyα radiation would be greatly absorbed by such interstellar medium. In fact, it was expected that only a few LAEs would be observed in the sky.

Neufeld (1990) also gave an analytic expression for the Lyα line profile emerging from an homogenius slab. This is,

$$J(\tau_0, x) = \frac{\sqrt{6}}{24} \frac{x^2}{a\tau_0\sqrt{\pi}} \left[\frac{1}{1 + \cosh\left(\frac{x^3}{a\tau_0} \sqrt{\frac{\pi^3}{54}} \right)} \right]. \tag{1.45}$$

In Fig.1.14 we show how the Lyα line profile of a slab varies with its column density. In general, the Lyα exhibits a symmetric double peak profile. The denser is the medium the more difficult is for Lyα photons to escape. In fact, for low $\mathrm{N_H}$ values the peak are close to Lyα, but as N_H grows, more scattering events take place and the peaks get away from Lyα.

1.4.3 The detection of the first LAEs

The first observations of LAEs were made in 1998.

- Steidel et al. (1998) studied a photometry selected sample of "Lyman break galaxies"[2] between redshift 2.2 and 3.4 to find overdesities of galaxies in the early Universe. Then, they performed spectroscopic observations of their targets to characterize the performance of their selection. Among other things, they found that a significant fraction of their targets exhibited a Lyα line. Moreover, they found that the Lyα line was slight more redshifted with respect to the galaxy spectrum.

- Cowie and Hu (1998) used a set of photometric bands to target LAEs at high redshift. They used 3 broad bands (B, V and I) and one narrow band filter centered at $\lambda = 5390$Å with FWHM$\sim$ 77Å. The broad bands main function was to detect the continuum of the galaxy, while the narrow band was designed to detect Lyα at redshift ~ 3.4. They made their LAE target selection by looking for objects with an excess in the narrow band. They found 12 strong Lyα emitters.

  > *"It is shown that there is a substantial population of galaxies at $z \sim 3.4$ which can be selected by Lyα emission."*
  >
  > Cowie and Hu (1998)

- Later, Hu et al. (1998) carried out spectroscopic observations of their previous LAE candidate sample. They found Lyα in a significant fraction of their targets. This work demonstrated the high performance of photometric narrow band observations to detect Lyα at high redshift. Additionally, they found that the Lyα line profiles were not symmetric.

From this point, many surveys targeting Lyα have been carried out. The first large LAE survey was the "Large Area Lyman Alpha Survey" (Rhoads et al. (2000) , Malhotra and Rhoads (2002)) where they covered an sky area of $36' \times 36'$ using narrow band observations. They found a great number of LAEs at $z \sim 4.5$.

---

[2] Lyman break galaxies constitute a galaxy population that is detected in the sky for their characteristic flux redaction at wavelengths smaller than Lyα.

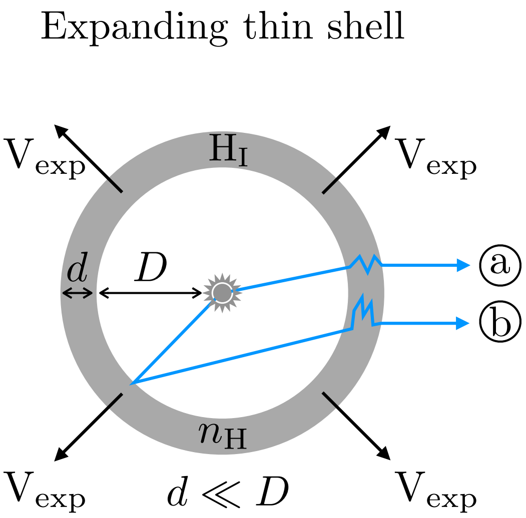


Fig. 1.15 Sketch of an expanding thin shell with a couple of photons as example. In grey we show the dense H_I region.

> *"... implying a net density of* ~ 4000 *Lyα emitters per square degree per unit redshift"*

Rhoads et al. (2000)

1.4.4 Outflows as an escape channel

By the year 2000, it had became obvious that LAEs were abundant in the high redshift Universe. After the prediction that millions of LAE could be found all over the sky the astrophysical community started wondering how Lyα photons could escape from galaxies. Neufeld (1990) estimations for the escape fraction were too low to explain the number of LAEs found.

One possible answer came from Neufeld (1991), where it was suggested that a multiphase intergalactic medium could enhance the escape of Lyα photons. In this scenario, instead of modeling the ISM as an homogeneous slab, it was composed by a optical thin medium and high density H_I clumps. Lyα would bounce in the surface of these H_I clouds without penetrating them, which would help their escape. However, this model could not explain the asymmetric Lyα profiles found in different observations. Note, that in the present, a multiphase intergalactic medium has been probed to reproduce many of the properties of LAEs (e.g. Gronke et al., 2016).

The solution at that time came form the hands of Ahn et al. (2000). In their work they found that the escape of Lyα photons through a expanding giant $\mathrm{H_I}$ bubble cloud explain the observed Lyα line profiles and facilitate the escape of Lyα photons. An illustration of their gas geometry (from now on Thin Shell) is shown in Fig.1.15. The Lyα source is in the center of an optically thin cavity (representing an $\mathrm{H_{II}}$ region). Photons travel freely through it until they reach a thin spherical layer of $\mathrm{H_I}$ that has a positive radial velocity $\mathrm{V_{exp}}$.

The approach to tackle the Lyα radiative transfer was innovative. Instead of solving the complex radiative transfer equations the authors designed a radiative transfer Monte Carlo code. In their code photons with random direction where emitted to the thin shell of $\mathrm{H_I}$. Then,

1. The optical depth of the cloud was computed along the direction of the photon.
2. The probability of interaction was computed and it was randomly decided if the photon interacted or not.
3. They computed the new direction of the photon and its new frequency.
4. These steps were repeated until the photon escaped.

They tracked the outcome of every interaction of photons with the neutral hydrogen. In this way, for each configuration of $\mathrm{N_H}$ and $\mathrm{V_{exp}}$ they obtained a different Lyα line profile.

This approach has proven to give the same solution as the analytic expression Neufeld (1990) when the gas geometry is an homogeneous slab. However, one of the disadvantages of this method is the computational time. A great number of random photons is necessary to obtain a line profile with good signal to noise. Even so, this approach is very useful and has been used and improved by the community (Ahn et al., 2000; Zheng and Miralda-Escudé, 2002; Ahn, 2003; Verhamme et al., 2006; Orsi et al., 2012; Gronke et al., 2016).

In Fig.1.16 we show a family of line profiles produced by a dustless Thin Shell at $T = 10^4 K$. These profiles were computed using `FLaREON` (see Chapter 3). In general the line profiles exhibit several peaks. However, in contrast with the homogeneous slab the relative intensity between the peaks is different. For example, at $\mathrm{V_{exp}} = 10\ km/s$ all the presented line profiles exhibit a peak bluer than Lyα and a peak redder. As the

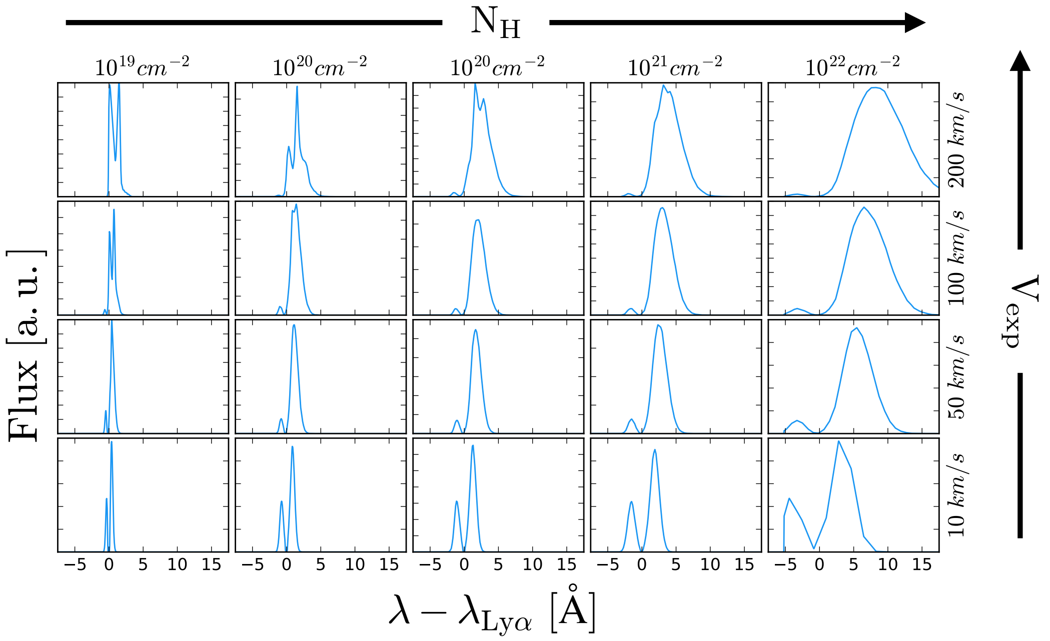


Fig. 1.16 Sample of Lyα line profiles that emerge of the thin shell geometry for several values of column density and macroscopic velocity.

outflow has an expanding bulk velocity it is more probable that Lyα interact with H_I atoms distancing from the source. As a consequence, normally the redder peak is more prominent. In fact, as V_{exp} gets higher, the blue peak contribution becomes negligible. Another clear effect is that the position and width of the peaks depend on N_H. In fact, Lyα lines become broader and more redshifted the higher the column density is.

Moreover, at very high values of V_{exp} several red peaks are produced. Theses peaks are produced by photons that instead of penetrating the thin shell (as photon 'a' of Fig.1.15) they bounce back inside the H_{II} cavity. Then they find another wall of the shell. The relative velocity between the point where the photon bounced the first time and the point of the second interaction is greater than V_{exp} and at maximum $2V_{exp}$ (without accounting for thermal motion). This causes that a second redder peak emerges at $V_{exp}= 200\ km/s$ and $N_H= 10^{20} cm^{-2}$. This kind of event is called *backscatter* (photon 'b' of Fig.1.15).

One of the most useful upgrades in the radiative transfer Monte Carlo code was the inclusion of dust. This allowed to study the escape fraction from these kind of outflows. In Fig.1.17 we show the escape fraction of Lyα photons as a function of the outflow properties (generated with `FLaREON`). As in the case of the homogeneous slab,

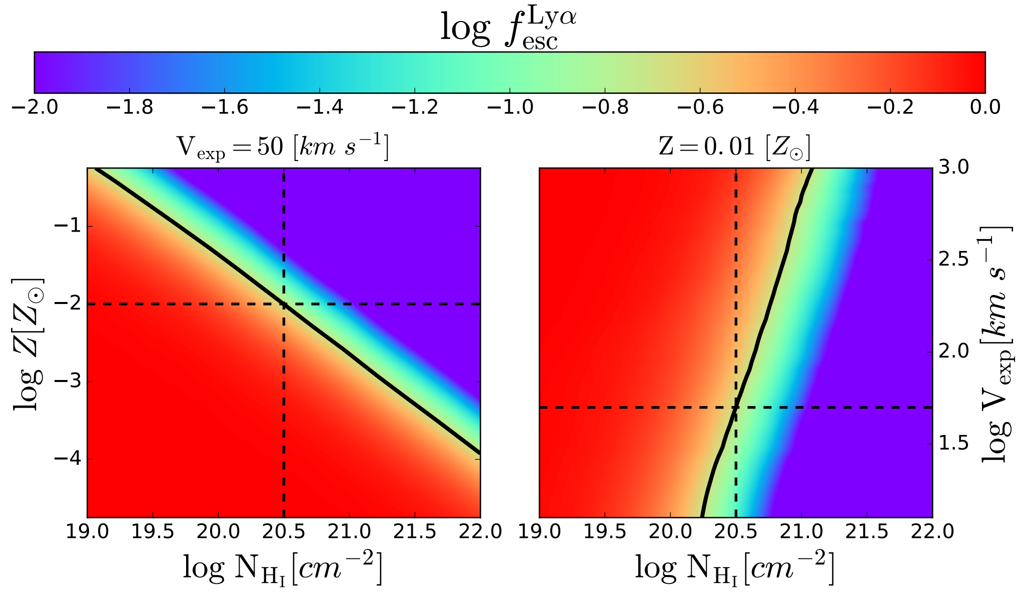


Fig. 1.17 Dependence of the Lyα escape fraction in the thin shell geometry. In the left we show the behaviour for a fixed value of $V_{\rm exp}$ =50 km/s. In the right panel we keep Z = 0.01 $Z_\odot$ fixed. In both panels the dark line shows $f_{\rm esc}^{\rm Ly\alpha} = 0.2$

the $f_{\rm esc}^{\rm Ly\alpha}$ decreases with metallicity and column density. However $f_{\rm esc}^{\rm Ly\alpha}$ is greater in the thin shell. For the typical values of $Z = 10^{-2} Z_\odot$ and $N_{\rm H} = 10^{20.5} cm^{-2}$, $f_{\rm esc}^{\rm Ly\alpha} \sim 0.01$ in the slab, while in the thin shell $f_{\rm esc}^{\rm Ly\alpha} \sim 0.2$. This enhancement of $f_{\rm esc}^{\rm Ly\alpha}$ is caused by the expansion velocity of the outflow. In fact, $f_{\rm esc}^{\rm Ly\alpha}$ increases rapidly with $V_{\rm exp}$. As the $H_{\rm I}$ atoms are distancing the Lyα source they received the photon redshifted (in average). In addition to the fact that the Lyα-$H_{\rm I}$ cross section peaks at Lyα and decreases really fast. This causes that the thin shell is more transparent than the static slab. Additionally, the backscatters help increasing $f_{\rm esc}^{\rm Ly\alpha}$. Thanks to the greater escape fraction of the thin shell and other outflow geometries later studied, the number counts of LAEs were explained (Garel et al., 2012; Orsi et al., 2012).

1.4.5 State-of-the-art of Lyman-α emitters

Half a century has passed since Partridge and Peebles (1967) suggested that Lyα could be observed in extragalactic sources. Nowadays, LAEs constitute one of the most popular galaxy population to study the early Universe. There are two main characteristics that makes Lyα such an interesting tool to trace galaxies at high redshift. First, Lyα falls in the optical range of the electromagnetic spectrum since when the Universe was 0.7 Gyr old ($z \sim 7.0$) until it was 3 Gyr old ($z \sim 2$). This allows ground experiments to be able to observed Ly$\alpha$when the Universe was young. The second feature is the high luminosity of Lyα in star forming galaxies. At the typical temperatures of the star formation regions, the probability of emitting Lyα in a H_I recombination event is the highest. Therefore, Lyα is the emission line with the highest intrinsic flux.

All over the world, scientific teams have made remarkable efforts to understand LAEs. On the theoretical side, the Lyα radiative transfer is a complex multi-scale problem. The journey of Lyα photons starts inside the star formation regions, which typical sizes about a few proper parsecs. Then, the Lyα photons interact with different galactic structures smaller than 1 pKpc (Behrens et al., 2019) and escape to the circumgalactic medium that can be as large as some hundreds of pKpc. Then, photons enter into the inter galactic medium, which interacts with Lyα along a few pMpc. Finally, in order to study cosmology with galaxies it is mandatory to explore a large enough volume (about a cGpc^3) to avoid cosmic variance. The multi-scale nature of the Lyα RT processes makes extremely difficult to properly study each of the different astrophysical environments at the same time. Indeed, the state-of-the-art cosmological simulations (e.g. Angulo et al., 2012; Nelson et al., 2015; Baugh et al., 2019) are still far from covering all this range of scales with high accuracy.

In order to make theoretical predictions about LAEs with the current tools, some authors have studied the LAEs as a galaxy population in a ΛCDM framework using radiative transfer Monte Carlo codes to simulate the Lyα RT inside the interstellar medium (e.g. Garel et al., 2012; Orsi et al., 2012). Other works have implemented the Lyα RT physics in *small* high-resolutions hydrodynamic simulations to study the escape of Lyα radiation from realistic ISM structure (e.g. Behrens et al., 2019). Meanwhile, the Lyα RT inside the ISM has been modeled in several ways. For example, Dijkstra et al. (2007) determine the IGM velocity field, density field and ionization state around galaxies with analytic expressions to determine the typical IGM absorption. Later, Laursen et al. (2011a) studied the IGM transmission in a cosmological hydrodynamic simulation by tracking Lyα rays along several lines of sight. Jeeson-Daniel et al. (2012) used a similar ray-tracking method to study the effects of the circumgalactic medium and IGM close to the galaxy during the epoch of reionization. Another technique to understand the Lyα RT in the IGM has been the direct implementation of radiative transfer Monte Carlo codes into cosmological hydrodynamic simulations (e.g. Zheng et al., 2010, 2011; Behrens et al., 2017).

On the observational side, a wide range of galaxy surveys have chased Lyman-α emitters across all the cosmic history. In the local Universe some authors have characterized the the *Blueberry galaxies*, local analogs of high redshift LAEs (Yang et al., 2017). At high redshift many galaxy surveys have studied this galaxy population to understand the mechanisms of galaxy formation evolution using photometric (e.g. Kashikawa et al., 2006; Ouchi et al., 2008; Guaita et al., 2010; Konno et al., 2016; Sobral et al., 2017a) and spectroscopic experiments (e.g. Esp; Herenz et al., 2017). An even during the epoch of reionization (Kashikawa et al., 2011; Sobral et al., 2015; Ouchi et al., 2018; Shibuya et al., 2018). Moreover, the scientific community will continue its efforts in the future by studying cosmology using LAEs as tracer in surveys such as HETDEX (Hill et al., 2008a), J-PAS (Benitez et al., 2014a) and DESI (DESI Collaboration et al., 2016).

1.5 Thesis outline

After this introduction to the Lyα radiative transfer and to the galaxy population known as Lyα emitters, we present the scientific results of this thesis. In the following subsections we present the abstracts from Chapter 2 to 6. Chapters 2, 3 and 4 are works submitted to journal, while Chapters 5 and 6 are still work in progress. Then, we make our conclusions and forecast future work in Chapter 7.

- Chapter 2: Lyα radiative transfer inside galaxies.

  Lyman-α emitters (LAEs) are a promising target to probe the large scale structure of the Universe at high redshifts, $z \gtrsim 2$. However, their detection is sensitive to radiative transfer effects that depend on local astrophysical conditions. Thus, modeling the bulk properties of this galaxy population remains challenging for theoretical models. Here we develop a physically-motivated scheme to predict LAEs in cosmological simulations. The escape of Lyα photons is computed using a Monte Carlo radiative transfer code which outputs a Lyα escape fraction. To speed-up the process of assigning escape fractions to individual galaxies, we employ fitting formulae that approximate the full Monte Carlo results within an accuracy of 10% for a broad range of column densities, gas metallicities and gas bulk velocities. We apply our methodology to the semi-analytical model `GALFORM` on a large N-body simulation. The Lyα photons escape through an outflowing neutral gas medium, implemented assuming different geometries. This results in different predictions for the typical column density and outflow velocities of the LAE population. To understand the impact of radiative transfer on our predictions, we contrast our models against a simple abundance matching assignment. Our full models populate LAEs in less massive haloes than what is obtained with abundance matching. Overall, radiative transfer effects result in better agreement when confronting the properties of LAEs against observational measurements. This suggest that incorporating the effects of Lyα radiative transfer in the analysis of this galaxy population, including their clustering, can be important for obtaining an unbiased interpretation of future data sets.

This chapter has been submitted and accepted in the astrophysical journal Monthly Notices of the Royal Astronomical Society with the ID MN-18-2387-MJ. The information of the article is: Monthly Notices of the Royal Astronomical Society, Volume 486, Issue 2, June 2019, Pages 1882–1906. This work can also be found at arXiv at `https://arxiv.org/abs/1807.00006`.

- Chapter 3: Fast computation of Lyα RT.

We present `FLaREON` (Fast Lyman-Alpha Radiative Escape from Outflowing Neutral gas), a public `Python` package that delivers fast and accurate Lyα escape fractions and line profiles over a wide range of outflow geometries and properties. The code incorporates different algorithms, such as interpolation and machine learning to predict Lyα line properties from a pre-computed grid of outflow configurations based on the outputs of a Monte Carlo radiative transfer code. Here we describe the algorithm, discuss its performance and illustrate some of its many applications. Most notably, `FLaREON` can be used to infer the physical properties of the outflowing medium from an observed Lyα line profile, including the escape fraction, or it can be run over millions of objects in a galaxy formation model to simulate the escape of Lyα photons in a cosmological volume.

This chapter has been submitted to the astrophysical journal Monthly Notices of the Royal Astronomical Society with the ID Draft MN-18-4256-MJ and it is pending of acceptance. This work can also be found in arXiv at `https://arxiv.org/abs/1811.09630`.

- Chapter 4: Lyα RT inside the intergalactic medium.

  In the near future galaxy surveys will target Lyman alpha emitting galaxies (LAEs) to unveil the nature of the dark energy. During the last years it has been suggested that the observability of LAEs is coupled to the large scale properties of the intergalactic medium. Such coupling could introduce distortions in the observed clustering of LAEs, adding a new potential difficulty to the interpretation of upcoming surveys. We present a model of LAEs that incorporates Lyα radiative transfer processes in the interstellar and intergalactic medium. The model is implemented on the `GALFORM` semi-analytic model of galaxy of formation and evolution. We find that the radiative transfer inside galaxies produces selection effects over galaxy properties. In particular, observed LAEs tend to have low metallicities and intermediate star formation rates. At low redshift we find no evidence of correlation between the spatial distribution of LAEs and the intergalactic medium properties. However, at high redshift the LAEs are linked to the line of sight velocity and density gradient of the intergalactic medium. The strength of the coupling depends on the outflow properties of the galaxies and redshift. This effect modifies the clustering of LAEs on large scales, adding non linear features. In particular, our model predicts modifications in the shape and position of the baryon acoustic oscillation peak. This work highlights the importance of including radiative transfer physics in the cosmological analysis of LAEs.

  This chapter has been submitted to the astrophysical journal Monthly Notices of the Royal Astronomical Society with the ID Draft MN-19-1264-MJ and it is pending of acceptance. This work can also be found at arXiv at `https://arxiv.org/abs/1904.04274`.

- Chapter 5: Imprints of the epoch of reionization.

  The epoch of reionization is nowadays one of the most interesting time windows in the evolution of the Universe. Before the formation of the first stars and galaxies the intergalactic medium (IGM) was in a neutral state. Then, the energetic photons emitted by galaxies progressively ionize the IGM. This transition is known as reionization. Here, we study the imprints that different histories of reionization leave in Lyα selected samples. We implement our LAE model introduced in previous chapters to this frame. For this goal, we mimic the typical topology of the local fraction of neutral hydrogen during the epoch of reionization. This divides the simulation volume in two kind of regions: H_{I} clumps and H_{II} bubbles. Lyman-α emitters are mainly observed in H_{II} regions, while they avoid H_{I} clouds. We find that the history of reionization affects different properties of Lyman-α emitters. For example the amplitude of their luminosity function anticorrelates with the global fraction of neutral hydrogen. Moreover, we find that the clustering of this galaxy population depends on the moment and duration of reionization in different ways. First, the reionization time modifies mainly the clustering amplitude, while the duration modifies the clustering shape. In this way, we also find very characteristics features in their clustering when it is divided in parallel and perpendicular along the line of sight. We conclude that Lyman-α emitters are a galaxy population that will help to determine the history of reionization with future surveys.

- Chapter 6: Chasing LAEs at redshift 2.2.

We present the first results of HiRULE: High Redshift Universe Lyman-α survEy, a program targeting LAEs and QSOs at $z \sim 2.2$ to study their cross-correlation over an unprecedented volume. In its initial phase, HiRULE consists in observing 20 deg^2 with the narrow band J0395 at T80Cam in the Observatiorio Astrofísico de Javalambre, allowing us to perform various cross-correlation studies of these two high-z populations. Key science cases include: a) Cross-correlation function of LAEs and QSOs, to quantify baryonic environmental effects acting at this redshift and affecting the small-scale clustering around QSOs; b) Lyα fluorescence around QSOs, from neutral gas surrounding QSOs; c) the properties of bright Ly-alpha emitters at $z \sim 2.2$, including their luminosity function and auto-correlation function. Here we present the data analysis and the measured LAE luminosity function.

Chapter 2

The impact of the interstellar medium in Lyα selected samples

This chapter has been published in Monthly Notices of the Royal Astronomical Society as Gurung-Lopez et al. 2019, Volume 486, Issue 2, June 2019, Pages 1882–1906

Resumen

Las galaxias emisoras de radiación Lyα (LAEs) son una población prometedora para trazar la estructura a gran escala del Universo a redshifts mayores que 2. Una de las principales ventajas de las LAEs es su alta luminosidad a la longitud de onda en reposo de Lyα, lo cual facilita su detección. Adicionalmente, debido la expansión de Hubble (Hubble, 1929), la linea Lyα es observable entre redshift 2 y 7, permitiendo que experimentos en tierra puedan medir estas galaxias. Sin embargo, su función de selección es compleja y depende de los procesos de transferencia radiactiva de Lyα, los cuales dependen de las condiciones astronómicas.

Presentamos un modelo teórico de LAEs basado en un código Monte Carlo de transferencia radiativa. Nuestro modelo puede ser aplicado a enormes volumes cosmológicos

debido a su eficiencia. En particular hemos aplicado nuestro modelo a la simulación de materia oscura `P-Millennium` y el modelo semianalítico de formación y evolución de galaxias `GALFORM` (Lacey et al., 2016).

Los códigos Monte Carlo de transferencia radiativa han demostrado ser una herramienta útil para entender cómo la radiación Lyα escapa de las galaxias. Desafortunadamente, los prohibitivos costes computacionales no permiten que estos códigos se corran en grandes volumes cosmológicos. Para evitar este problema hemos desarrollado unas expresiones analíticas para computar la fracción de fotones Lyα que escapan ($f_{\rm esc}^{\rm Ly\alpha}$). Estas expresiones son precisas en un gran rango de velocidad de expansión $\rm V_{exp}$, densidad de columna de hidrógeno neutro $\rm N_H$ y metalicidades Z.

Siguiendo nuestra metodología, computamos $f_{\rm esc}^{\rm Ly\alpha}$ para cada galaxia en función de $\rm V_{exp}$,$\rm N_H$ and Z. Estas cantidades son calculadas usando las propiedades de cada galaxia, como su tamaño, tasa de formación estelar o masa de halo. Los parámetros libres que regulan el calculo de estas cantidades son calibrados para reproducir las funciones de luminosidad observadas de LAEs. Después de la calibración encontramos que las tres geométricas estudiadas en este capitulo pueden reproducir las funciones de luminosidad observadas a bajo corrimiento al rojo. Sin embargo, solo el thin shell y el wind consiguen hacerlo a alto redshift, donde el bicono falla. Concluimos que la geometría biconica (tal y como es presentada aquí) a alto desplazamiento al rojo es menos favorecida que las otras. En el siguiente capitulo rediseñamos esta geometría modelándola más realisticamente haciendo que las propiedades dependan del ángulo de visión.

Hemos analizado la abundancia relativa de LAEs dividiendo las funciones de luminosidad en función de diversas propiedades. La masa del halo o la masa estelar no están significativamente correlacionadas con la luminosidad de Lyα. Las funciones de luminosidad tienden a estar dominadas por galaxias de baja masa. Sin embargo, cuando las funciones de luminosidad son divididas en tasa de formación estelar encontramos una correlación positiva con la luminosidad de Lyα. Finalmente, cuando la función de luminosidad es dividida en bines de metalicidad encontramos bastante dispersión para $\log Z < 2$. Es mas, la contribución de altas metalicidades ($\log Z > 2$) a la parte luminosa es pequeña.

También comparamos las propiedades de muestras seleccionadas por su emisión en Lyα con las propiedades del resto de galaxias. Encontramos que los LAEs tienden a habitar halos de materia oscura relativamente poco masivos. Adicionalmente, las galaxias con los episodios de formación estelar más fuertes no son seleccionadas cómo LAE, ya que las metalicidades de este tipo de galaxia hacen que la fracción de escape sea pequeña.

Para validar nuestras predicciones de $f_{\rm esc}^{\rm Ly\alpha}$, hemos comparado nuestras muestras de LAEs con las observaciones llevadas a cabo por Oyarzún et al. (2017). Encontramos que nuestro modelo reproduce muy bien la relación observada entre $f_{\rm esc}^{\rm Ly\alpha}$ y la tasa de formación estelar. Las muestras de LAE que incluyen transferencia radiativa reproducen con éxito esta anti-correlación y su dispersión. Sin embargo, la relación $f_{\rm esc}^{\rm Ly\alpha}$-$M_*$ esta desplazada 0.5 *dex* in M_* con respecto a Oyarzún et al. (2017). Esta diferencia puede estar causada por los diferentes modelos de síntesis de formación estelar usados por `GALFORM` y Oyarzún et al. (2017). Finalmente, encontramos que nuestra muestra de LAE basada en una relación monofónica entre la tasa de formación estelar y la luminosidad de Lyα (es decir, obviando los procesos de transferencia radiativa) no reproduce las tendencias observadas. Esto subraya el papel crucial que juega la transferencia radiativa.

También estudiamos la población de halos de materia oscura que habitan las LAEs en nuestros modelos. Encontramos diferencias entre las muestras que incluyen transferencia radiativa y los que la excluyen. A bajo corrimiento al rojo, los modelos con transferencia radiativa predicen que las LAE habitan halos de menor masa y por lo tanto su función de correlación presenta menos amplitud. Esta tendencia se invierte a alto desplazamiento al rojo.

Finalmente, comparamos la distribución espacial de las LAEs en nuestros modelos con observaciones. A desplazamiento al rojo 2.2, 3.0 y 5.7 la amplitud de la función de correlación reproduce correctamente las observaciones. Sin embargo a corrimiento al rojo 6.7 la amplitud no concuerda. Aunque el efecto del medio intergaláctico no esta muy claro, trabajos previos (Zheng et al., 2011) indican que este podría tener un gran efecto en muestras seleccionadas por su luminosidad en Lyα.

En este capitulo demostramos la importancia de la transferencia radiativa determinando la función de selección de las LAEs basada en propiedades galácticas como la metalicidad, la masa o la tasa de formación estelar. Por una parte, las tendencias observadas no se puede reproducir simplemente con una relación monotonía entre la tasa de formación estelar y luminosidad Lyα. Por otra parte, La inclusión de la transferencia raditiva cambia las propiedades de las LAEs observadas. Todo esto hace de extrema importancia la inclusión de la transferencia radiativa en los modelos para entender la formación y evolución de galaxias.

Abstract

Lyman-α emitters (LAEs) are a promising target to probe the large scale structure of the Universe at high redshifts, $z \gtrsim 2$. However, their detection is sensitive to radiative transfer effects that depend on local astrophysical conditions. Thus, modeling the bulk properties of this galaxy population remains challenging for theoretical models. Here we develop a physically-motivated scheme to predict LAEs in cosmological simulations. The escape of Lyα photons is computed using a Monte Carlo radiative transfer code which outputs a Lyα escape fraction. To speed-up the process of assigning escape fractions to individual galaxies, we employ fitting formulae that approximate the full Monte Carlo results within an accuracy of 10% for a broad range of column densities, gas metallicities and gas bulk velocities. We apply our methodology to the semi-analytical model `GALFORM` on a large N-body simulation. The Lyα photons escape through an outflowing neutral gas medium, implemented assuming different geometries. This results in different predictions for the typical column density and outflow velocities of the LAE population. To understand the impact of radiative transfer on our predictions, we contrast our models against a simple abundance matching assignment. Our full models populate LAEs in less massive haloes than what is obtained with abundance matching. Overall, radiative transfer effects result in better agreement when confronting the properties of LAEs against observational measurements. This suggest that incorporating the effects of Lyα radiative transfer in the analysis of this galaxy population, including their clustering, can be important for obtaining an unbiased interpretation of future datasets.

2.1 Introduction

During the past two decades, surveys targeting the Lyα emission in star-forming galaxies, the so-called Lyα emitters (LAEs), have detected objects out to redshift $z \sim 7$ (e.g. Steidel et al., 1996; Hu et al., 1998; Rhoads et al., 2000; Malhotra and Rhoads, 2002; Taniguchi et al., 2005; Kashikawa et al., 2006; Guaita et al., 2010; Konno et al., 2016; Sobral et al., 2017a).The study of this galaxy population has allowed us to explore the kinematics of the interstellar medium (ISM) in high redshift galaxies (Shapley et al., 2003; Steidel et al., 2010, 2011; Kulas et al., 2011; Guaita et al., 2017; Chisholm et al., 2017), the large scale structure (Gawiser et al., 2007; Orsi et al., 2008; Ouchi et al., 2010; Bielby et al., 2016; Kusakabe et al., 2018; Ouchi et al., 2018), the epoch of reionization (Santos et al., 2004; Kashikawa et al., 2006; Dayal et al., 2011; Inoue et al., 2018) and to test galaxy formation models (Le Delliou et al., 2006; Kobayashi et al., 2007; Nagamine et al., 2010; Orsi et al., 2012).

Despite the success in detecting progressively larger samples of LAEs, their physical interpretation has proven to be a difficult challenge (see Dijkstra, 2017, for a review). Lyα photons are easily scattered by neutral hydrogen, causing a large increase in the path that the photon needs to travel through neutral hydrogen clouds (e.g. Harrington, 1973; Neufeld, 1990). This results in an increased probability of interaction with dust grains, and thus, absorption. Hence, the Lyα radiative transfer through a neutral medium reduces the Lyα flux that escapes the galaxy and also modifies the line profile, since each scattering event changes the frequency of the photons. These physical processes also take place in the surrounding intergalactic medium (IGM) of galaxies and can also modify the observed Lyα flux and line profile (Santos et al., 2004; Dijkstra et al., 2011).

Analytical approximations for Lyα radiative transfer have been derived for over-simplistic neutral gas configurations (e.g. Harrington, 1973; Neufeld, 1990; Dijkstra et al., 2006). More realistic configurations can be explored with a Monte Carlo algorithm. Individual Lyα photons are generated inside a neutral hydrogen cloud with a given geometry, kinematics and temperature. The path of Lyα photons is tracked including their interactions, which produce scattering events, until the photons escape or are absorbed by dust. This approach has been studied in several scenarios (Ahn et al., 2000; Zheng and Miralda-Escudé, 2002; Ahn, 2003; Verhamme et al., 2006; Gronke et al., 2016). Most notably, Monte Carlo radiative transfer has shown to reproduce the diversity of observed Lyα line profiles by allowing photons to escape through an outflowing medium (e.g. Schaerer and Verhamme, 2008; Orsi et al., 2012).

Theoretical models of galaxy formation have introduced the effect of radiative transfer in different approximate ways to predict the properties of the LAE population. The first model of LAEs in a hierarchical galaxy formation framework implemented a constant escape fraction of Lyα photons to reproduce their observed abundance and clustering (Le Delliou et al., 2005, 2006; Orsi et al., 2008). Further attempts introduced radiative transfer effects over simple geometries in semi-analytical models (Orsi et al., 2012; Garel et al., 2012). Cosmological hydrodynamical simulations also incorporated Lyα radiative transfer in post-processing. One approach has been to track Lyα rays to simulate different lines of sight (e.g. Laursen and Sommer-Larsen, 2007; Laursen et al., 2009, 2011b) over small volumes. With a Monte Carlo radiative transfer code, Zheng et al. (2010) showed that the proper treatment of Lyα photons radiative transfer has dramatic effects on the clustering of LAEs. However, recently, Behrens et al. (2017) found no significant change in the clustering of LAEs after implementing Lyα radiative transfer in the Illustris simulation (Nelson et al., 2015), and attribute the claims of Zheng et al. (2010) about the clustering of LAEs to resolution effects.

In the next years many ground-based large surveys such as HETDEX (Hill et al., 2008b), J-PAS (Benitez et al., 2014b) and space missions like ATLAS-Probe (Wang et al., 2018a), will aim to detect LAEs over large areas to trace the large scale structure (LSS) at high redshifts. Such measurements could potentially deliver cosmological constraints in redshift ranges well above those currently targeted by Multi-Object Spectroscopic surveys. With progressively larger and more accurate datasets, it becomes crucial to improve our theoretical understanding of galaxies as tracers of the underlying matter distribution (Orsi and Angulo, 2018). One of our aims in this work is to understand the impact of radiative transfer effects on clustering measurements.

The model for the Lyα luminosity of star-forming galaxies presented here is based on a fast implementation of a Monte Carlo radiative transfer. To avoid the prohibitively long time that it would take to run a Monte Carlo code over millions of galaxies, we develop fitting formulae that reproduce the full Monte Carlo results accurately. To illustrate the potential of our model, we apply this methodology to the semi-analytic model `GALFORM` run over an N-body simulation. This is a first paper in a series that explores the properties of galaxies selected by their Lyα luminosity. Here we focus on the impact of the Lyα RT in defining the properties of the LAE galaxy population. In a forthcoming paper we implement the impact of the intergalactic medium (IGM) and the effects of reionization on the LAE population.

The outline of this paper is as follows: in §2 we develop fitting formulae to predict the escape fraction of Lyman alpha photons through outflows. In §3, we describe our model for LAEs that combines galaxy formation physics and Lyα radiative transfer in addition to the implementation of the Lyα RT in a galaxy formation model is presented. We analyze the LAE population predicted by our model in §4. We discuss our results in §5. Finally, conclusions and future work are summarized in §6.

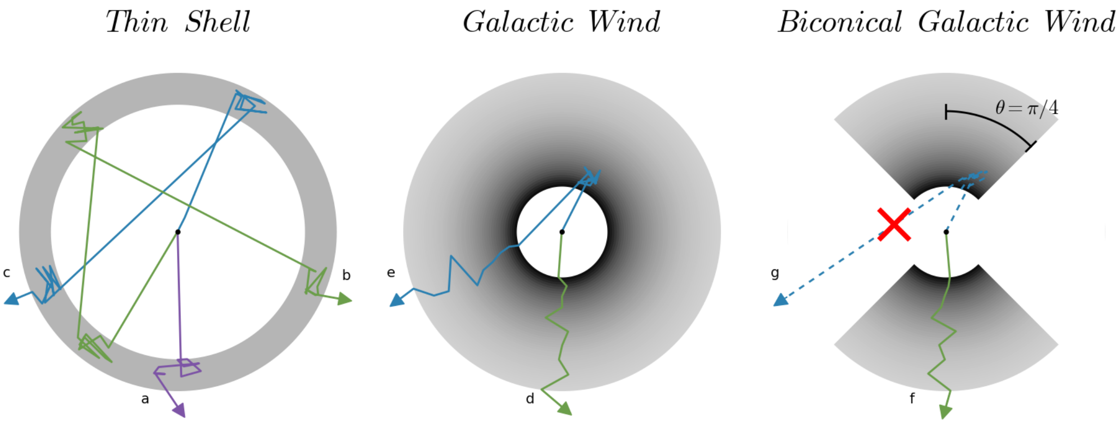


Fig. 2.1 A schematic illustration of the different outflow geometries implemented in this work: *Thin Shell* (left), *Wind* (middle) and *Biconical Wind* (right). The gas density is represented by the gray colour scale. Different possible trajectories of photons are labeled from a to g. The red cross over photon g illustrates the point where this photon is absorbed by the medium.

2.2 Model ingredients

In this section we describe our model ingredients and the methodology we follow to predict the properties of LAEs in a cosmological simulation.

2.2.1 Lyα radiative transfer

We track the scattering, absorption and escape of Lyα photons making use of the Monte Carlo radiative transfer code described in Orsi et al. (2012), which has been made publicly available [1]. This code is similar to others in the literature (e.g. Zheng and Miralda-Escudé, 2002; Ahn, 2003, 2004; Dijkstra et al., 2006; Verhamme et al., 2006; Laursen and Sommer-Larsen, 2007; Barnes and Haehnelt, 2010, and references therein). A detailed review of Lyα radiative transfer can be found in Dijkstra (2017). Below we summarize the main features of the Orsi et al. (2012) code that are most relevant to this work.

The code receives as input a configuration of a 3D neutral gas geometry, temperature, expansion velocity $V_{\rm exp}$, neutral hydrogen column density $N_{\rm H}$ and optical depth of dust τ_a. In this work we use a monochromatic source of photons, i.e., all photons

---

[1] https://github.com/aaorsi/LyaRT

are generated at Lyman α wavelength, $\lambda_0 = 1215.68$Å. For a given gas distribution, the code generates a Lyα photon with a random direction and follows its interactions with hydrogen and dust until it is either absorbed by dust or escapes from the neutral gas medium. Every interaction with a hydrogen atom results in a scattering event that changes the direction and frequency of the photon. Interactions with dust, on the other hand, can change the direction of the photon or result in absorption depending on the assumed albedo of the dust grains. The process is repeated for N_p photons, recording in the end the frequency of every photon that escaped and those that were absorbed by dust grains. This allows us to compute the escape fraction $f_{\rm esc}^{\rm Ly\alpha}$ and wavelength distribution (i.e. the Lyα line profile) for every outflow geometry over which both the neutral gas and the dust are distributed. In this work we implemented three different outflow geometries, which are illustrated in Fig. 2.1.

1. *Thin Shell.* This geometry consists of an expanding isothermal homogeneous spherical shell similar to ones found in literature (e.g. Verhamme et al., 2006; Garel et al., 2012; Orsi et al., 2012) . This spherical shell is thin and can be described by an inner and an outer radius, $R_{\rm in}$ and $R_{\rm out}$ respectively, which satisfy $\rm R_{in}/R_{out} = 0.9$. The shell is expanding outwards, thus it has a radial macroscopic velocity $\rm V_{exp} > 0$. The neutral hydrogen column density is given by:

$$N_H = \frac{M_H}{4\pi m_H R_{\rm out}^2}, \tag{2.1}$$

   where M_H is the total neutral hydrogen mass and m_H is the mass of a hydrogen atom.

   The empty cavity in the center of the shell produces photon *backscatterings*, i.e. photons can bounce back into the empty cavity multiple times, as illustrated by photons b and c in Fig. 2.1.

2. *Galactic Wind.* This geometry consists of an expanding spherical gas distribution with a central empty cavity of radius $R_{\rm Wind}$ (Orsi et al., 2012). The gas is isothermal and is expanding radially at a constant velocity $V_{\rm exp}$. Unlike the *Thin Shell*, the gas is distributed with a radial density profile given by:

$$\rho_{\rm H}({\rm r}) = \begin{cases} 0 & r < R_{\rm Wind} \\ \frac{\dot{M}_H}{4\pi m_H r^2 V_{\rm exp}} & r \geqslant R_{\rm Wind}, \end{cases} \tag{2.2}$$

where $\dot{M}_H$ is the ejection neutral hydrogen mass rate. Thus, the column density in the *Wind* geometry is

$$N_H = \frac{\dot{M}_H}{4\pi m_H R_{\rm Wind} V_{\rm exp}}. \tag{2.3}$$

This geometry is illustrated in the middle panel of Fig. 2.1.

We define a large outer radius $R_{\rm out} = 20 R_{\rm Wind}$ where the computation is forced to end and any photon that have reached this radius is considered to have escaped. We have checked that for greater values of $R_{\rm out}$ the code provides the same line profile and escape fraction. Thus, we conclude that our results converge for our choice of $R_{\rm out}$.

3. *Biconical Wind*. This geometry shares the same properties of the *Wind*, but additionally it features an aperture angle, $\theta_{\rm cone}$, which defines the volume of gas and dust similar to Zheng and Wallace (2014). In particular we arbitrarily set $\theta_{\rm cone} = \pi/4$. The resulting polar asymmetry is thus the main difference between the two previous geometries and this one. We checked that there is no strong dependence of $f_{\rm esc}^{\rm Ly\alpha}$ with the line of sight within the outflow aperture. This geometry is shown in the right panel of Fig. 2.1.

In the biconical geometry we force photons to be emitted from the center of the geometry (as in the other geometries) and within the aperture of the bicone, i.e. no photons are emitted outside the bicone. Additionally, due to the empty regions in this geometry, photons that scatter off the internal cavity and escape off the bicone are considered absorbed by the external medium (e.g. photon g in Fig. 2.1). This is equivalent to assuming that there is a dusty optically thick medium surrounding the bicone.

Furthermore, in this geometry we define the escape fraction of Lyα photons as the ratio between the number of photons emitted towards the bicone and the

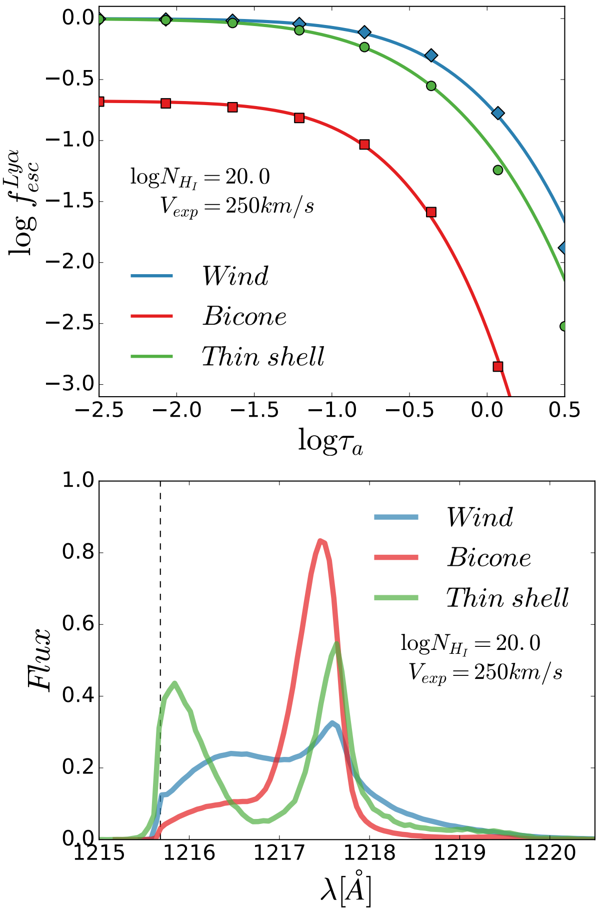


Fig. 2.2 (Top) $f_{esc}^{Ly\alpha}$ versus the dust optical depth τ_a for different geometries in outflows with the same physical properties ($V_{\rm exp}$ and $N_{\rm H}$), as indicated in the figures. The output of the radiative transfer code is represented by green circles, blue diamonds and red squares for the *Thin Shell*, galactic wind and biconical geometries respectively. Additionally, our analytical fit is represented by solid lines with the same color code as the code's output. (Bottom) Lyα line profile for different geometries with the same physical properties. In colored lines the radiative transfer code output is plotted for the *Thin Shell* geometry (green), the galactic wind (blue) and the biconical galactic wind (red).

number of photons that escaped through the bicone. This implies that in every galaxy the outflow geometry is pointing towards the observer. We note that the actual fraction of galaxies with a bicone orientated towards the observer would be $\Delta\gamma_b/4\pi$, where $\Delta\gamma_b$ is solid angle. In fact, if we take into account the galaxy orientation in our models we find that for the bicone the observed number counts of LAEs is too low (as we discuss later). Therefore, our biconical model represents an upper limit. For further details on the bicone implementation see §2.7.1.

Fig. 2.2 illustrates the difference between the Lyα escape fraction (left panel) and line profile (right panel) predicted by each geometry, for a particular choice of column density and expansion velocity. As expected, the escape fraction decreases towards higher values of τ_a in all geometries, as greater amounts of dust absorb more photons. However, the impact on the geometry of the medium is evident: even if the three configurations have the same $N_{\rm H}$ and $V_{\rm exp}$, photons have the highest escape fractions from the *Wind* geometry, and the lowest from the *Bicone*. This is due to the complicated Lyα RT. For example, as in the *Bicone* configuration photons that leak through the empty cavity are considered absorbed, the escape fraction does not reach 1 even if there is no dust in the outflow, making a great difference with respect to the other two geometries. Additionally, even if the *Wind* and *Thin Shell* configurations share spherical symmetry (unlike the *Bicone*) the dependence of $f_{\rm esc}^{\rm Ly\alpha}$ on $N_{\rm H}$, $V_{\rm exp}$ and τ_a is different due to the distinctive hydrogen density radial profiles of the two configurations. This dependence on the geometry does not only affect the $f_{\rm esc}^{\rm Ly\alpha}$ but also the line profile of the Lyα emission. The predicted line shape changes dramatically from a geometry to another: in the case of the *Wind* it is a broad line, for the *Bicone* it is a narrow line and for the *Thin Shell* it assumes a double-peak profile. We use these three different outflow geometries to estimate the variance in the LAEs population depending on the geometry.

2.2.2 Fitting formulae for Lyα radiative transfer

As discussed in §2.2.1, the Monte Carlo radiative transfer code can take a long time to run for a given configuration of parameters. For a single photon, the average number of scatterings, and thus, calculations, grows as a power-law function of the column density of the medium (Harrington, 1973). In the parameter space explored here, the completion time of the code can vary from a few seconds up to a few hours in the most extreme cases. Applying this directly in a catalog of millions of objects would result in prohibitively long execution times.

To overcome this, we develop empirical (measured from the radiative transfer Monte Carlo code) expressions that approximate the results of the Monte Carlo runs. We start by constructing a grid to scan the parameter space with ~ 450 configurations spanning the ranges $10 \leq {\rm V_{exp}[km\ s^{-1}]} \leq 1000$ and $-2.5 \leq \log \tau_a \leq 0.5$. For the

Table 2.1 Constant parameter values used to derive the escape fraction of the different geometries

Thin Shell	
$k_1 = k_{11} V_{\rm exp}^{k_{12}}$	$k_2 = k_{21} V_{\rm exp}^{k_{22}}$
$k_{11} = k_{111} (\log N_{H_{18}})^2 + k_{112} \log N_{H_{18}} + k_{113}$	$k_{21} = 10^{-0.0368}$
$k_{12} = k_{121} (\log N_{H_{18}})^3 + k_{122} (\log N_{H_{18}})^2 + k_{123} \log N_{H_{18}} + k_{124}$	$k_{22} = 10^{-1.556}$
$k_{111} = 10^{2.109}$	
$k_{112} = -10^{2.745}$	
$k_{113} = 10^{2.954}$	
$k_{121} = 10^{-1.785}$	
$k_{122} = -10^{-0.730}$	
$k_{123} = 10^{-0.155}$	
$k_{124} = -10^{0.151}$	
Galactic Wind	
$k_1 = k_{11} V_{\rm exp}^{k_{12}}$	$k_2 = k_{21} V_{\rm exp}^{k_{22}}$
$k_{11} = k_{111} N_{H_{18}}^{k_{112}}$	$k_{21} = 10^{0.0137}$
$k_{12} = k_{121} (\log N_{H_{18}})^2 + k_{122} \log N_{H_{18}}$	$k_{22} = 10^{-1.62}$
$k_{111} = 10^{0.471}$	
$k_{112} = 10^{-0.244}$	
$k_{121} = 10^{-1.82}$	
$k_{122} = -10^{-0.667}$	
Biconical Wind	
$k_1 = k_{11} V_{\rm exp}^{k_{12}} + k_{13}$	$k_2 = k_{21} V_{\rm exp}^{k_{22}}$
$k_{11} = 10^{3.229}$	$k_{21} = 10^{0.0470}$
$k_{12} = -10^{-0.0752}$	$k_{22} = 10^{-1.490}$
$k_{13} = k_{131}\ N_{H_{18}}^{k_{132}} + k_{133}$	
$k_{131} = 10^{-0.580}$	
$k_{132} = 10^{-0.238}$	
$k_{133} = 10^{0.700}$	
$k_3 = k_{31} V_{\rm exp}^{k_{32}} + k_{33}$	
$k_{31} = 10^{k_{311} (\log N_{H_{18}})^2 + k_{312}}$	
$k_{32} = k_{321} (\log N_{H_{18}})^2 + k_{322} \log N_{H_{18}} + k_{323}$	
$k_{33} = 10^{-0.0779}$	
$k_{311} = 10^{-0.874}$	$k_{321} = -10^{-1.226}$
$k_{312} = 10^{0.571}$	$k_{322} = 10^{-0.477}$
	$k_{323} = -10^{0.292}$

Thin Shell and *Wind* geometries we covered $19.0 \leq \log(\mathrm{N_H}[\mathrm{cm}^{-2}]) \leq 22.5$. However, due to the excessive computational time in the *Bicone* we restrict our analysis to $19.0 \leq \log(\mathrm{N_H}[\mathrm{cm}^{-2}]) \leq 22.0$. We run the Monte Carlo code with 10^4 photons and obtain the Lyα escape fraction, $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ as a function of τ_a , N_H and, V_{exp}. We set the gas temperature to 10^4K.

To construct an analytic expression for $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ we start from a generalized form of the expression for the $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ in a homogeneous, static slab derived in Neufeld (1990):

$$f_{\mathrm{esc}}^{\mathrm{an}} = k_3 \left[\cosh\sqrt{k_1 \tau_a^{k_2}}\right]^{-1} , \tag{2.4}$$

where k_1 and k_2 are functions of N_H and V_{exp} for all geometries. Additionally, k_3 is set to 1 in the *Thin Shell* and *Wind* geometries, but is a function, $k_3(\mathrm{N_H}, \mathrm{V_{exp}}) < 1$, in the *Bicone*, since, in this geometry, the escape fraction is always less than 1 (see section 2.2.1). We perform a Monte Carlo Markov Chain (MCMC) with the `emcee`[2] code (Foreman-Mackey et al., 2013a) to determine the functional form of k_1, k_2 and k_3, by minimizing the function

$$\chi^2 = \sum_{N_H, V_{\mathrm{exp}}, \tau_a} \left(\frac{f_{\mathrm{esc}}^{\mathrm{MC}} - f_{\mathrm{esc}}^{\mathrm{an}}}{\sigma_{\mathrm{MC}}}\right)^2 , \tag{2.5}$$

where f_{esc}^{MC} corresponds to the escape fraction of photons obtained with the MC code over each configuration in the grid, and σ_{MC} is the error in the calculation of the escape fraction, given by the dispersion in a binomial distribution with probability of success f_{esc}^{MC} :

$$\sigma_{MC} = z_{1-\alpha/2}\sqrt{\frac{f_{\mathrm{esc}}^{\mathrm{MC}}(1 - f_{\mathrm{esc}}^{\mathrm{MC}})}{N}}, \tag{2.6}$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$-th percentile of the standard normal distribution. In particular we use the quantile 95, i.e. $\alpha = 0.1$. Additionally, N is the number of generated photons in each configuration.

The functional form and parameter values of the fits for $k_1(N_H, V_{\mathrm{exp}})$, $k_2(N_H, V_{\mathrm{exp}})$ and $k_3(N_H, V_{\mathrm{exp}})$ for each geometry are shown in Table 2.1.

Fig. 2.2 compares the $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ computed analytically with Eq. (3.6) and with the free parameters obtained with the MCMC (lines), and that obtained with the full MC RT

---

[2] http://dfm.io/emcee/current/

Table 2.2 Free parameters as defined in equations 2.7 and 2.8 after the calibration with the observed luminosity function for different geometries and redshifts.

redshift	Geometry	$\log \kappa_{V,disk}$	$\log \kappa_{V,bulge}$	$\log \kappa_{N,disk}$	$\log \kappa_{N,bulge}$
$z = 2.2$	Thin Shell	4.440	4.911	-12.367	-11.839
	Wind	4.857	4.914	-7.065	-5.338
	Bicone	4.982	4.258	-8.140	-7.249
$z = 3.0$	Thin Shell	4.337	4.549	-12.465	-11.915
	Wind	4.691	4.769	-7.440	-5.166
	Bicone	4.896	4.338	-8.436	-6.404
$z = 5.7$	Thin Shell	4.737	4.428	-13.906	-11.808
	Wind	4.660	3.782	-8.292	-6.180
	Bicone	4.612	3.590	-8.078	-7.614
$z = 6.7$	Thin Shell	4.659	4.279	-13.81	-11.934
	Wind	4.589	3.871	-8.073	-5.910
	Bicone	4.455	3.561	-7.848	-7.647

code (symbols) for a given values of $\mathrm{N_H}$ and $\mathrm{V_{exp}}$ and the three different geometries. The analytical expression reproduce remarkably well the results of the full MC RT code.

The accuracy of our analytic expressions varies with $\tau_{\rm a}$, $\mathrm{V_{exp}}$, $\mathrm{N_H}$ and the geometry. In particular, there is a strong dependence on $\tau_{\rm a}$: for every geometry we find that the accuracy decreases with increasing $\tau_{\rm a}$. We find that, in general, the discrepancy with the full MC RT code in configurations with $\tau_{\rm a} > 10^{-0.5}$ becomes greater than 10%. Galaxies with such a large dust absorption, in general, will not be observed as a LAE so we are not concerned about the low accuracy at high $\tau_{\rm a}$. Additionally, we checked that, after calibration of our LAEs model (see §2.3.1), less than 2% of the Lyα selected galaxies (LAEs) in each geometry have $\tau_{\rm a} > 10^{-0.5}$, making the contribution of these galaxies negligible. For galaxies with $10^{-1.5} < \tau_{\rm a} < 10^{-0.5}$, the discrepancy is just a few percents for $\mathrm{N_H}$ between 10^{19} and $10^{22.5}\mathrm{cm^{-2}}$ and $\mathrm{V_{exp}}$ between 80 and 1000 km s^{-1}. Moreover, for $\tau_{\rm a} < 10^{-1.5}$ the discrepancy is typically below the 1% in the same parameter range. Appendix 4.9.1 show the performance of our analytic expressions and the $\tau_{\rm a}$, $\mathrm{V_{exp}}$ and $\mathrm{N_H}$ distributions of the LAE samples.

2.2.3 Simulation and semi-analytical model.

We combine the radiative transfer code described above with the semi-analytical model of galaxy formation `GALFORM` (Lacey et al., 2016) run on the `P-Millennium` N-body simulation (Baugh et al., in prep.).

The `P-Millennium` is a state-of-the-art dark matter only N-body simulation using the Plank cosmology: $H_0 = 67.77$ km s^{-1}Mpc^{-1}, $\Omega_\Lambda = 0.693$, $\Omega_M = 0.307$, $\sigma_8 = 0.8288$ (Planck Collaboration et al., 2016). The box size is 542.16 cMpc h^{-1} and the particle mass $M_p = 1.061 \times 10^8$ $M_\odot$ h^{-1} (5040^3 dark matter particles). Between the initial redshift, $z = 127$, and the present, $z = 0$, there are 272 snapshots. In this work we use snapshots 77, 84, 120 and 136 corresponding to redshifts 6.7, 5.7, 3.0, 2.2, respectively.

A full review on semi-analytical models of galaxy formation can be found in Baugh (2006). The variant of `GALFORM` used in this work is based on earlier versions described in Cole et al. (2000); Baugh et al. (2005) and Bower et al. (2006). In brief, `GALFORM` computes the properties of the galaxy population following the hierarchical growth of dark matter halos. Halo merger trees are extracted from an N-body simulation (the `P-Millennium` in our case), so the model can also predict the spatial distribution and peculiar velocities of galaxies.

In `GALFORM`, galaxies are formed and evolve as a result of the following processes: i) the radiative cooling and the shock-heating of gas inside halos; ii) the subsequent cooling of gas forming a disk at the bottom of the potential well; iii) quiescent star formation in the disk and starbursts in bulges resulting from disk instabilities and galaxy mergers; iv) feedback processes (supernovae, AGN and photoionization) regulating the star formation, and v) the chemical enrichment of stars and gas that results from star-formation and feedback episodes. Additionally, the variant of `GALFORM` used in this work assumes different initial mass functions (IMFs). In particular, a Kennicutt (1983) IMF is used for quiescent star formation, while for starburst modes a top-heavy IMF is implemented (see Lacey et al., 2016, for more details).

`GALFORM` generates a composite spectral energy distribution (SED) for each individual galaxy based on its star-forming history and computes the rate of emission of hydrogen ionizing photons, $\dot{Q}_H$, by integrating the galaxy SED over wavelengths bluer than the Lyman break at $\lambda = 912$Å. All ionizing photons are assumed to be absorbed by the neutral medium. Then case B recombination (Osterbrock, 1989) is used to compute the intrinsic line luminosity of Lyα, where a fraction of 0.66 of ionizing photons contribute to generating Lyα photons.

In this work we only make use of `GALFORM` galaxies with $M_{\rm stellar} > 10^7 {\rm M}_\odot h^{-1}$. This stellar mass threshold translates in dark matter halos with mass $\rm M_h > 10^{10} M_\odot h^{-1}$, well above the halo mass resolution limit of the simulation, $\sim 2\ 10^9 {\rm M}_\odot h^{-1}$.

Radiative transfer parameters

To combine the Monte Carlo radiative transfer code with `GALFORM`, we need to derive the parameters that define the neutral gas configuration from the galaxy output properties. In particular, the column density $N_{\rm H}$, expansion velocity $V_{\rm exp}$ and the optical depth of dust $\tau_{\rm a}$ are key to determine the escape fraction. Motivated by observational works (e.g. Cazzoli et al., 2016), the expansion velocity is computed for the three geometries as:

$$\rm V_{exp,c} = \kappa_{\mathit{V,c}} SFR_{\mathit{c}} \frac{\mathit{r_c}}{\mathit{M_*}}, \tag{2.7}$$

where the index c denotes the galaxy component (disk or bulge), SFR_c and r_c are the SFR[$\rm M_\odot$/Gyr/h] and half mass radius [pMpc] of each galaxy component, $M_*[M_\odot/h]$ is the total stellar mass of the galaxy and $\kappa_{\rm V,c}$ are two (one per galaxy component) free parameters. Note that, in our model, the photons generated in each galaxy component do not interact with the outflow corresponding to the other galaxy component.

The neutral hydrogen column density is computed in different ways depending on the geometry (see section 2.2.1) :

$$\rm N_{H,c} = \begin{cases} \kappa_{\mathit{N,c}} \frac{\mathit{M}_{cold,c}}{\mathit{r_c^2}} & \mathit{Thin\ Shell} \\ \kappa_{\mathit{N,c}} \frac{\mathit{M}_{cold,c}}{\mathit{r_c} \mathit{V}_{exp,c}} & \mathit{Wind} \text{ and } \mathit{Bicone} \end{cases} \tag{2.8}$$

where $M_{\rm cold,c}$ and $\kappa_{\rm N,c}$ are, respectively, the cold gas mass and a free parameter of the galaxy component c.

All the free parameters linking `GALFORM` properties to $\rm V_{exp}$ and $\rm N_H$ are calibrated by fitting the observed LAE luminosity function at different redshifts. For further details see §2.3.1.

Finally, the τ_a is computed for every geometry as:

$$\tau_{a,c} = (1 - A_{\mathrm{Ly}\alpha})\frac{E_\odot}{Z_\odot}N_{H,c}Z_c, \tag{2.9}$$

where $A_{Ly\alpha} = 0.39$ is the albedo at the Lyα wavelength, $E_\odot = 1.77 \times 10^{-21}\mathrm{cm}^{-2}$ is the ratio τ_a/N_H for solar metallicity, $Z_\odot = 0.02$ (Granato et al., 2000) and Z_c is the cold gas metallicity of the galaxy component c.

The intrinsic Lyα LF predicted by `GALFORM` (see Figure 2.3) results from two populations: normal star forming galaxies (populating the low luminosity range) and galaxies with an ongoing star formation burst (populating the high luminosity range). Consequently, the values of $\kappa_{N,disk}$ and $\kappa_{V,disk}$ control the shape of the faint-end LF, whereas $\kappa_{N,bulge}$ and $\kappa_{V,bulge}$ control the bright end of the LF. In both regimes, increasing (decreasing) $\kappa_{N,c}$ leads to an increase (decrease) of the N_H distribution. This leads to a decreasing (increasing) in the resulting $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ distribution and thus lowers (increases) the number of galaxies with higher luminosities. Also, increasing (decreasing) $\kappa_{V,c}$ leads to a increase (decrease) of the $\mathrm{V}_{\mathrm{exp}}$ distribution, increasing (decreasing) $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ and the number of galaxies with high luminosities.

2.3 Implementing Lyα radiative transfer in a semi-analytical model.

In this section we describe how we incorporate the Lyα radiative transfer processes inside the semi-analytical galaxies from `GALFORM`. We make use of the fitting formula described above to predict the Lyα escape fraction and line profiles. The strategy to fit the value of the free parameters of Eqs. 2.7 and 2.8 is described below.

2.3.1 Calibrating the model.

In order to calibrate the model and compute the values of the free parameters for each geometry, we fit our model to the observed LAE luminosity function at redshifts $z = 2.2, 3, 5.7$ and 6.7. We run `emcee` (Foreman-Mackey et al., 2013b) to perform an MCMC to find the values of $\kappa_{\mathrm{N,c}}, \kappa_{\mathrm{V,c}}$. The dynamical range of each free parameter is determined by limiting the expansion velocity and column densities of each component

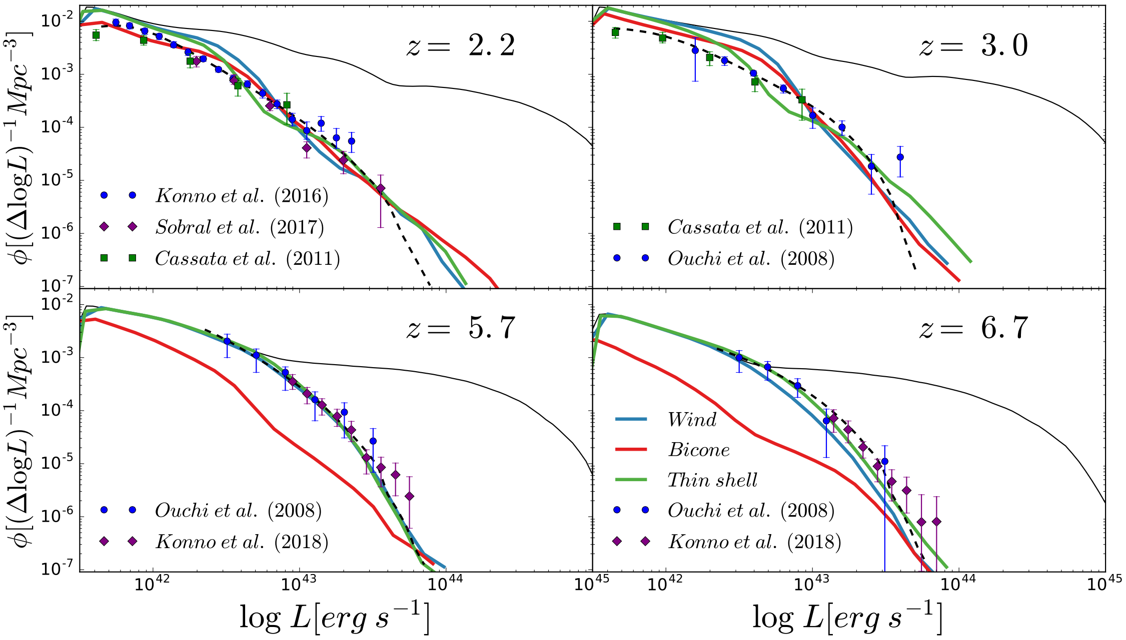


Fig. 2.3 LAE LF at redshift 2.2 (top left), 3.0, (top right), 5.7 (bottom left) and 6.7 (bottom right). The LF computed for different geometries is plotted as colored continuum lines, in blue for the *Wind* geometry, in red for the *Bicone* geometry and in green the *Thin Shell* geometry. In continuum black we show the intrinsic Lyα LF. The black dashed lines show the combined LF that is fitted that, at the same time, is the AM-noRT LF (detailed in §2.4) LF. At redshift 2.2 we also show the LF observed by Kono et al 2016 (blue dots), Sobral et al. 2016 (purple diamonds) and Cassata et al 2011 (green squares). At redshift 3.0 we show the LF observed by Cassata et al 2015 (green squares) and Ouchi et al. 2008 (blue dots). At redshift 5.7 and 6.7 we show the LF observed by Ouchi et al. 2008 (blue dots) and Konno et al. 2018 (purple diamonds).

to lie within $80 < V_{\rm exp}[\rm km\ s^{-1}] < 1000$ and $19.0 < \log(N_H[\rm cm^{-2}]) < 22.5$ for at least 90% of the resulting galaxy population with Lyα rest frame equivalent width $\rm EW_0 > 20$ Å and Lyα luminosity $\rm L_{Ly\alpha} > 10^{41.5}erg\ s^{-1}$. These limits are imposed by the range of validity of the fitting formulae to derive the escape fraction (see §2.2.2).

This calibration is done independently for each outflow geometry and individual redshift bin. To combine multiple observed LFs at redshift 2.2 and 3.0 we compute a 5th-order polynomial fit (in logarithm of Lyα luminosity - logarithm LF space) taking into account the uncertainties of each survey to obtain a single curve that represents the observational measurements. We choose to use a 5th-order polynomial at these redshifts as some recent works suggest that the typical Schechter function is not able to reproduce the observe LF (Konno et al., 2016; Sobral et al., 2017a). Additionally, at redshift 5.7 and 6.7 we use the best fitting Schechter function to the observed LAE LF computed by Konno et al. (2018). The LF used to calibrate our model are shown in Fig.2.3 in black dashed lines.

The model Lyα luminosity of galaxies, for each geometry and choice of [$\kappa_{V,disk}$, $\kappa_{V,bulge}$, $\kappa_{N,disk}$, $\kappa_{N,bulge}$] is computed as follows: i) we compute the intrinsic Lyα luminosity of each component, $L^{0}_{\mathrm{Ly}\alpha}$, of each galaxy, which is directly proportional to the ionizing photon production $\dot{Q}_H$ predicted by `GALFORM`; ii) we compute $V_{exp,disk}$, $N_{H_I,disk}$ and $\tau_{a,disk}$ using Eqs. (2.8) and (4.3); iii) we obtain $f^{\mathrm{Ly}\alpha}_{\mathrm{esc}}$ for each galaxy component using Eq. (3.6); iv) the observed Lyα luminosity of each component is obtained by multiplying the intrinsic luminosities by their respective $f^{\mathrm{Ly}\alpha}_{\mathrm{esc}}$; and v) the total Ly$\alpha$ luminosity for each galaxy is the sum of the observed luminosity of each component (disk + bulge).

Fig. 2.3 shows the observed LAE LF (points), the full `GALFORM` intrinsic Lyα LF (thin black line), the predictions for each geometry (thick colored lines) using the free parameters that result from the MCMC (listed in table 4.2) at the different redshifts implemented in this work.

The intrinsic Lyα LF in divided into two populations: normal SFR galaxies in the low luminosity range and starburst galaxies in the high luminosity range. In general, in `GALFORM` the galaxy disk component in dominated by a quiescent SFR while in bulges the main mode of star formation is starburst, although quiescent star formation is also included. Additionally, in `GALFORM` the quiescent SFR and the starburst have different IMFs, which produces the bumps in the LF. On one hand, at lower redshifts, the predicted intrinsic LF is above the observations at all luminosities, thus galaxies at these redshifts require a significant $f^{\mathrm{Ly}\alpha}_{\mathrm{esc}} < 1$ in order to reduce the amplitude of the LF. On the other hand, at redshifts 5.7 and 6.7, the intrinsic LF at low $\mathrm{L}_{\mathrm{Ly}\alpha}$ (disk-dominated region) matches observations, implying that galaxies in this range must have $f^{\mathrm{Ly}\alpha}_{\mathrm{esc}} \sim 1$. Additionally, the intrinsic high redshift LF at high luminosities (bulge-dominated regime) requires $f^{\mathrm{Ly}\alpha}_{\mathrm{esc}} < 1$.

In general, the MCMC approach finds good matching solutions for the models including the Lyα radiative transfer. First, we find that the *Thin Shell* is consistent with the measured LF at at all redshifts. Secondly, the *Wind* geometry performs quite well at $z = 2.2, 3.0$ and 5.7 while at $z = 6.7$ it underpredicts the number density of LAE. However, we have checked that by allowing $V_{\rm exp}$ to be slightly higher, the observed LF is matched at redshift 6.7 as well. In the third place, the *Bicone* geometry matches the observed LF at $z = 2.2$ and 3.0 while at $z = 5.7$ and 6.7 it fails. The low abundance of LAEs predicted with the *Bicone* geometry arises due to the low escape fractions predicted by this geometry. In fact, at high redshifts, faint Lyα emitters require escape fractions close to 1 to match the observed LFs, and this is not possible in the *Bicone* geometry by construction, as shown in Fig. 2.2.

2.3.2 A simplified model with no Lyα radiative transfer

In order to highlight how radiative transfer changes the properties of LAEs, we compare the properties of our model with an abundance matching approach. We perform a simple SFR-Lyα mapping where no Lyα radiative transfer is taken into account. We refer to this model variant as 'AM-noRT'.

To construct the AM-noRT model, we rank galaxies by their SFR. We assign a Lyα luminosity to each galaxy based on their total SFR in a monotonic way. Objects with the highest SFR are assigned the brightest Lyα luminosity. We compute the Lyα equivalent width using the assigned Lyα luminosity and continuum luminosity around the Lyα frequency provided by `GALFORM`. Lyα luminosities are assigned recursively towards lower luminosities such that the Lyα observed luminosity function (using the EW_0 cut of each survey) is recovered at each redshift. The resulting Lyα luminosity distribution is shown in Fig. 2.3 as dashed black line. We compute a $f_{\rm esc}^{\rm Ly\alpha}$, which corresponds to the ratio between the assigned $Ly\alpha$ luminosity and the intrinsic one.

In contrast with our RT models, the $f_{\rm esc}^{\rm Ly\alpha}$ in the SFR-only model does not depend on properties such as the cold gas mass or the galaxy metallicity. Due to the way that Lyα luminosities are computed, the resulting $f_{\rm esc}^{\rm Ly\alpha}$ can be higher than 1 in some cases.

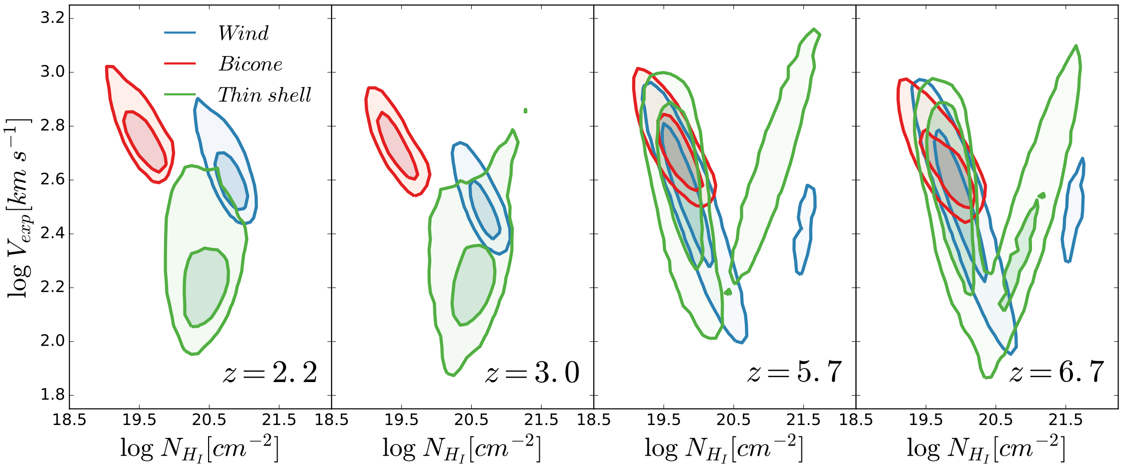


Fig. 2.4 Ouflow expansion velocity and neutral hydrogen column density distributions for each redshift ($z = 2.2, 3.0, 5.7$ and 6.7 from left to right) and for each geometry color coded as stated in the legend. The dark and light shaded contours enclose the 40 and 80 percentiles of the galaxy population, respectively.

2.4 Results.

In this section we describe the main predictions of our radiative transfer model when applied to `GALFORM` with the different outflow geometries.

2.4.1 The $\mathrm{N_H}$ and $\mathrm{V_{exp}}$ distributions.

Since the parameters in our model are calibrated to match the observed LFs for each geometry independently, the resulting distributions of $\mathrm{N_H}$ and $\mathrm{V_{exp}}$ are different for each configuration. Though this work unless it is different stated, we define LAE as a galaxy with a Lyα restframe equivalent width $\mathrm{EW_0} > 20$Å as typically in the literature (e.g. Ouchi et al., 2018). In this section, we rank LAEs by their luminosity and select the brightest to achieve a number density of $10^{-3}\mathrm{h^3cMpc^{-3}}$.

Fig. 2.4 shows the distribution of V_{exp} and $\mathrm{N_H}$ for each geometry. Since each quantity is computed for the disk and bulge component of each galaxy separately, we weight each component by their observed Lyα luminosity to build the distributions shown in Fig. 2.4.

Overall, the $\mathrm{V_{exp}} - \mathrm{N_H}$ distribution is relatively compact at redshifts (z=2.2,3.0) and more extended at higher redshift (z=5.7,6.7). The *Thin Shell* tends to have lower V_{exp} and $\mathrm{N_H}$ than the *Wind* geometry. Additionally, there is a strong difference between low and high redshift for these two distributions, while, in the case of the *Bicone*, remains generally unchanged across cosmic time. Additionally, most of the galaxies lie within the $f_{\rm esc}^{\rm Ly\alpha}$ analytic expression optimal accuracy region defined in §3. Moreover, we have checked that the fraction of galaxies outside the this region is lower than a 7% for every geometry and redshift.

Typical values the V_{exp} are found to be around 150km/s and 300km/s for the *Thin Shell* and *Wind* geometries respectively at $z = 2.2, 3.0$. Meanwhile, $\rm N_H$ is found at $\sim 10^{20.5}\rm cm^{-2}$ for the *Thin Shell* and $\sim 10^{20.8}\rm cm^{-2}$ for the *Wind.* Notably, at higher redshifts, z = 5.7 and 6.7, the distributions acquire a 'V' shape (especially visible for the *Thin Shell*) due to the division of each `GALFORM` galaxy into a disk and bulge and the significant difference in $f_{\rm esc}^{\rm Ly\alpha}$ for starburst and normal SFR galaxies at these redshifts. Lower column densities are favored by disk-dominated galaxies, requiring a higher $f_{\rm esc}^{\rm Ly\alpha}$ in order to fit the LF. the distribution of these galaxies peak around $\rm N_H \sim 10^{19.7}cm^{-2}$ and $\rm V_{exp} \sim 300km/s$. Bulge-dominates starbursts require a lower $f_{\rm esc}^{\rm Ly\alpha}$ to fit the LF, thus they favor high N_H and low $V_{\rm exp}$ distributions centered around $10^{21.0}\rm cm^{-2}$ and 200km/s respectively.

The *Bicone* geometry displays noticeable differences with respect to the other two geometries. The *Bicone* $\rm V_{exp} - N_H$ distributions are very similar across the different redshifts used in this work and present the available highest $\rm V_{exp}$ and lowest $\rm N_H$ distributions (peaking around 600km/s and $10^{19.2}\rm cm^{-2}$ respectively), maximizing as much as possible the escape of Lyα photons. This is due to the fact that the typical $f_{\rm esc}^{\rm Ly\alpha}$ is always lower in the *Bicone* compared to the other geometries, and it never reaches 1.

On one hand, at low redshift (z = 2.2, 3.0) our models fit the observed Lyα LF. At these redshifts, the *Bicone* is able to predict low enough $\rm N_H$ and high enough $\rm V_{exp}$ to reproduce the observed LAE number counts. The $\rm N_H$ and $\rm V_{exp}$ values in the *Bicone* are lower and higher respectively than in the other geometries, thus enhancing $f_{\rm esc}^{\rm Ly\alpha}$.

On the other hand, at high redshift (z = 5.7, 6.7) the *Bicone* is not able to select lower (higher) $\rm N_H$ ($\rm V_{exp}$) as these values fall outside of the parameter grid (see §2.2.2). Thus, the Lyα LF with this geometry is not able to fit the observed LF, as shown above.

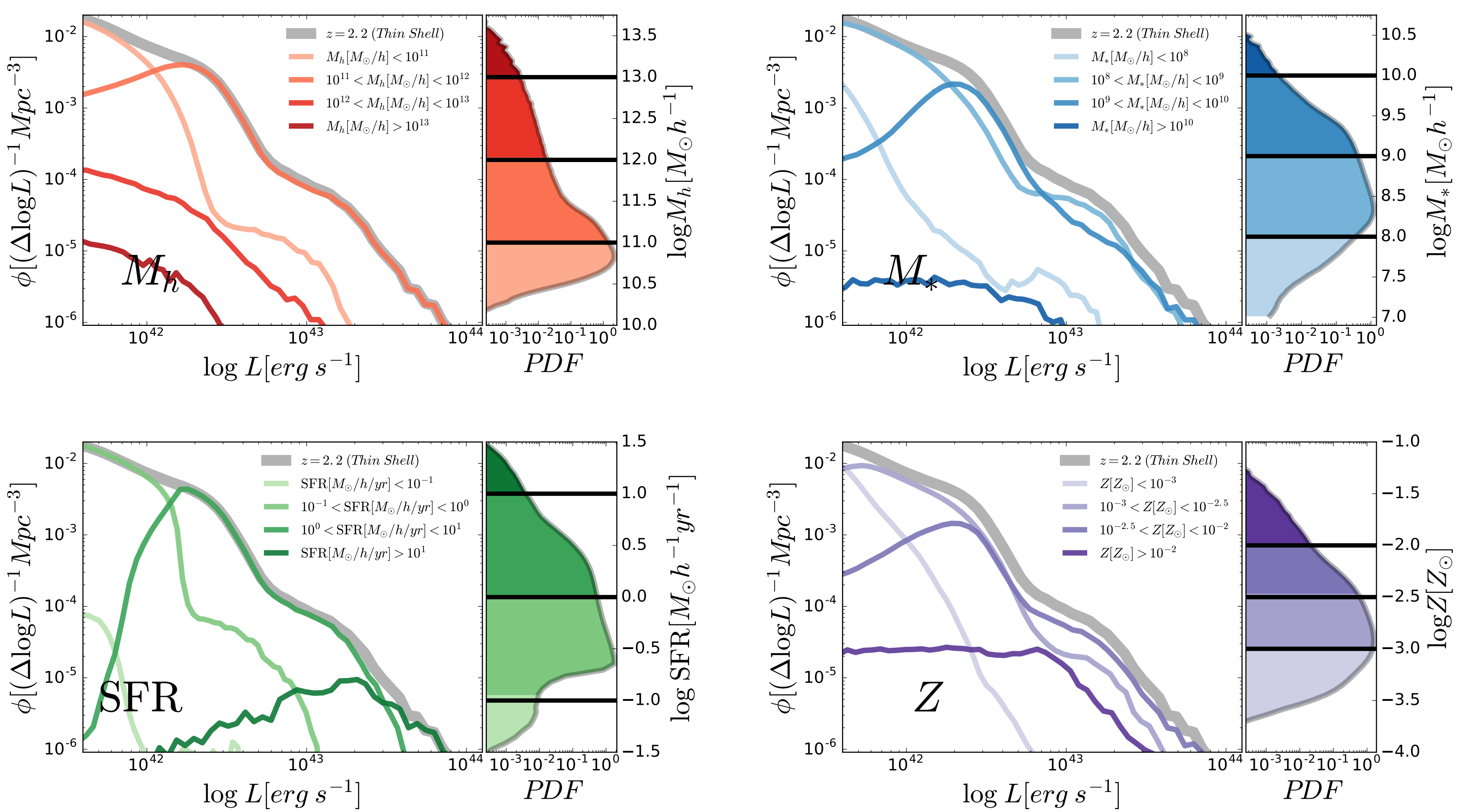


Fig. 2.5 **Left panels** : Break down of the Lyα LF at $z = 2.2$ of galaxies with Lyα EW$_0 > 20$Å in bins of halo mass (upper left), stellar mass (upper right), star formation rate (lower left) and metallicity (lower right). The bins are indicated in the legends. In each quantity the bins are represented in lighter colors for low values and darker as they increase. The total LF is plotted in thick gray line. **Right panels** : The probability distribution function of the different properties. In black we show the bin cuts.

2.4.2 Breaking down the Lyα LF

To illustrate the properties of LAEs, Fig.4.9 shows the Lyα LF obtained with the *Thin Shell* geometry at $z = 3.0$, split by the contribution of different ranges of halo and stellar mass, star formation rate and gas metallicity. We note that other redshifts and geometries show a similar behavior to what is shown in Fig. 4.9. Here we are analyzing a subsample composed of every LAE ($\mathrm{EW}_0 > 20$Å) with Lyα luminosity $> 10^{41.5} erg\ s^{-1}$.

When splitting the LF based on the halo mass of LAEs (upper-left panel), we find that the majority of LAEs are hosted by haloes of moderate mass, $M_{\mathrm{halo}} \sim 10^{11-12}\mathrm{M}_{\odot}h^{-1}$ which dominates the bright and moderate luminosities. LAEs with host halo masses below $M_{\mathrm{halo}} \lesssim 10^{11}\mathrm{M}_{\odot}h^{-1}$ dominate the very faint end of the LF, with $\mathrm{L}_{\mathrm{Ly}\alpha} \approx 10^{41}\mathrm{erg\ s}^{-1}$. Finally, the most massive haloes host galaxies do not contribute significantly to the LF shape. Furthermore, we have checked that there is no clear correlation between halo mass and Lyα luminosity.

In the upper right panel in Fig. 4.9 the LF is split according to the stellar mass of the emitting galaxy. The whole body of the LF is dominated by LAEs with stellar mass about $M_{\mathrm{stellar}} \sim 10^{8-10}\mathrm{M}_{\odot}h^{-1}$. Moreover, galaxies with a very low ($M_{\mathrm{stellar}} < 10^{8}\mathrm{M}_{\odot}h^{-1}$) or a very high ($M_{\mathrm{stellar}} > 10^{9}\mathrm{M}_{\odot}h^{-1}$) stellar mass do not contribute to bright or the faint ends. As in the M_{h} case, we do not find any clear correlation between stellar mass and Lyα luminosity.

The star formation rate, as expected, contributes in a roughly monotonic way to the Lyα LF. The faint-end of the Lyα LF is dominated by galaxies with low $\log(\mathrm{SFR}[M_{\odot}/h/yr]) \sim -0.5$. Additionally, the intermediate luminosities are dominated by moderate SFR $\sim 1 - 10[M_{\odot}/h/yr]$ while the bright end is populated by galaxies with the highest SFR (although with a significant scatter). Note that this trend only means that the $\mathrm{L}_{\mathrm{Ly}\alpha}$ of LAEs scales with SFR, but not that every galaxy with high SFR would result in a LAE. Finally, we note that typically, galaxies with SFR $< 0.1[M_{\odot}/h/yr]$ do not contribute to the LF in our luminosity ranges, although they might dominate the very faint end of the Lyα LF.

The break down of the Lyα LF in terms of gas metallicity is less intuitive. Naively one would expect to find an anti-correlation between metallicity and Lyα luminosity, since $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ decreases with increasing dust, and thus, metallicity. However, we find the

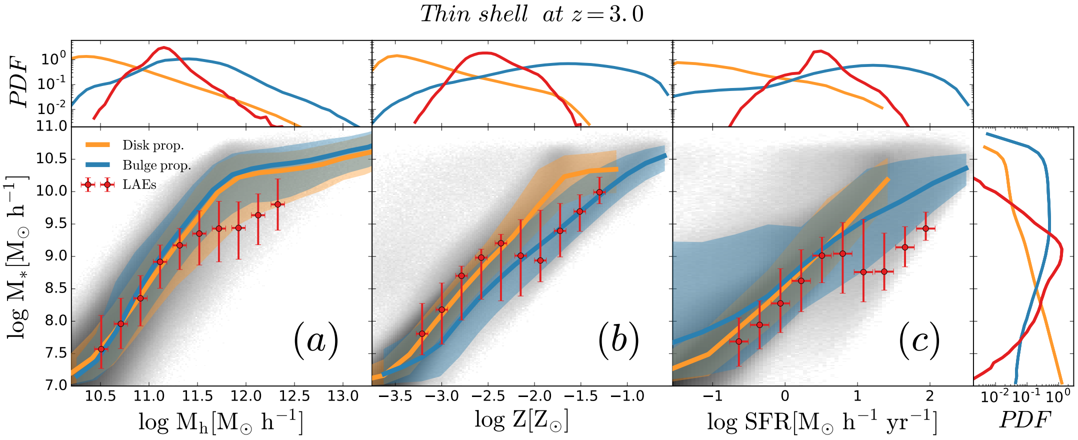


Fig. 2.6 **a)** The stellar mass - halo mass distribution at $z = 3.0$. The gray shaded region shows the distribution for the full `GALFORM` sample. The solid yellow and blue lines and correspond to the median of `GALFORM` central galaxies disk and bulges properties respectively. The shade regions show the 10-90 percentiles. The red dots show the *Thin Shell* LAE sample median, 10-90 percentiles (vertical) and the bin size (horizontal). **b)** Same as a) but for the stellar mass - metallicity distribution. **c)** Same as a) but for the stellar mass - star formation distribution. The top panels show the distributions of the halo mass, star formation and metallicity, respectively, for the full `GALFORM` (yellow and blue for disk and bulge dominated respectively) and the *Thin Shell* model (red). The stellar mass distribution is shown in the right vertical panel.

opposite: for LAEs with $\log(Z) < -2$, the low metallicity bins contribute to the lower luminosities and vice versa. This trend is broken for $\log(Z) > -2$ due to the low $f_{\rm esc}^{\rm Ly\alpha}$ at this metallicity range. The galaxies with highest Z do not contribute anymore to the bright end but to low and average luminosities . This leads to the bulk of the Lyα emitter population being dominated by galaxies with average metallicities, spanning the range $-3 < \log(Z) < -2$. We dig deeper in this relation in §2.4.3.

2.4.3 The bulk properties of LAEs.

In this section we analyze the galaxy properties of our simulated LAE, focusing on the results at redshift $z = 2.2$ and for the *Thin Shell* geometry (we checked that different geometries and redshifts give similar results). We restrict our analysis only to the brightest central LAEs with a 10^{-3}cMpc$^{-3}h^3$ number density cut (we check that different number density cuts produce similar results), and we compare it with the properties of the underlying population of central galaxies, i.e., the full population of galaxies predicted by `GALFORM` with $M_{\rm stellar} > 10^7 {\rm M}_\odot h^{-1}$.

Figure 2.6 shows some physical properties of the LAEs (red dots) and for the general population of galaxies from `GALFORM` selected using the same number density cut as the LAEs (yellow for disk properties and blue for bulge properties). Each panel includes the distribution of halo mass M_h, star formation rate SFR, metallicity Z and stellar mass M_* and the correlation between $M_* - M_h$, $M_* - {\rm SFR}$ and $M_* - Z$.

The M_h distribution in the LAE sample peaks at intermediate $M_h \sim 10^{11} M_\odot h^{-1}$ and spans between $10^{10.5} - 10^{12} [M_\odot h^{-1}]$. LAEs halos trace the massive end of the disk-dominated $\rm M_h$ distribution while avoiding the most massive dark matter halos, even if they host the strongest starburst episodes. This is caused by the SFR $- Z$ predicted by `GALFORM` that associates high metallicites (low $f_{\rm esc}^{\rm Ly\alpha}$) to high SFR.

The metallicity and the SFR of the LAE sample behave in a similar way due to the tight SFR $- Z$ relation. The bulk of the LAE sample peaks at intermediate values of Z and SFR, avoiding the extremes of the full `GALFORM` distribution. In particular, the galaxies with the highest SFR are not selected as LAE as the metallicity is also too high, causing a lower $f_{\rm esc}^{Ly\alpha}$. Additionally, the galaxies with extreme low Z are not selected either as their SFR in too low in these galaxies.

The $M_* - M_h$ relations (Fig. 2.6) for disk and bulge-dominated galaxies behaves in the same way. On the other hand, in the LAE sample this relation is the same as in the underlying galaxy population up to the peak of the $\rm M_h$ and $\rm M_*$ distributions, where the relation flattens for higher halo masses. In the high halo mass regime, LAEs typically have lower stellar masses than the overall average. This behavior is given by the tight SFR $-$ Z relation causing $f_{\rm esc}^{Ly\alpha}$ to be lower for galaxies with higher $\rm M_*$ as they become more dust rich.

In the LAE sample, the SFR $- Z$ relation is consistent with the bulk of the disk-dominated galaxies for $Z < 10^{-2.5} Z_\odot$. After a transition around $Z \sim 10^{-2.2}$, $Z_\odot$ is consistent with starburst galaxies. At metallicities below that transition the LAE SFR $-$ Z relation is slightly above the overall relation.

In the LAE sample the $M_* - {\rm SFR}$ relation is below the full `GALFORM` relation. This implies that for a fixed stellar mass, galaxies with higher SFR are selected, as the intrinsic $L_{Ly\alpha}$ correlates directly with the SFR.

2.4.4 The predicted Lyα $f_{\rm esc}^{\rm Ly\alpha}$ against observational estimates

In this section we compare our model predictions for the $f_{\rm esc}^{\rm Ly\alpha}$ against observational estimates from Oyarzún et al. (2017) at $z = 3$. In order to mimic their sample selection function we select galaxies with $10^{7.6}M_\odot < M_* < 10^{10.6}M_\odot$, and $L_{\rm Ly\alpha} > 10^{41.5}$erg s^{-1}.

Fig. 2.7 shows the relation between the Lyα $f_{\rm esc}^{\rm Ly\alpha}$ and the SFR and stellar mass. The $f_{\rm esc}^{\rm Ly\alpha}$ computed in Oyarzún et al. (2017) displays a noticeable anti-correlation between SFR and $f_{\rm esc}^{\rm Ly\alpha}$. In the models including RT galaxies with higher SFR have lower values of $f_{\rm esc}^{\rm Ly\alpha}$, in remarkable agreement with the observational estimates. The scatter in the observational data of Oyarzún et al. (2017) is consistent with the spread predicted by our models. This anti-correlation is caused by intrinsic link between SFR and Z. Even if the $V_{\rm exp}$ is higher for greater SFR (equation 2.7), dust plays the major role in the escape of Lyα photons and reduces $f_{\rm esc}^{\rm Ly\alpha}$.

The stellar mass is also anti-correlated with the $f_{\rm esc}^{\rm Ly\alpha}$, as shown in the right panel of Fig. 2.7. This is due to the known correlation between M_* and Z. Although our models reproduce the observationally inferred trend, the stellar masses predicted by `GALFORM` are systematically larger by $\sim$ 0.5dex. In particular, Mitchell et al. (2013) compared stellar masses using different stellar population models and `GALFORM` stellar masses. They conclude that the uncertainty in the choice of stellar population model and parameters can lead to biased stellar masses by a factor of 0.5dex.

Interestingly, the abundance matching model AM-noRT does not display the same trends found in Oyarzún et al. (2017), highlighting the importance of considering radiative transfer effects to predict LAE galaxy properties consistent with observational datasets.

2.4.5 The dark matter haloes hosting LAEs

In the following we study the properties of dark matter halos hosting LAEs. To compare different model predictions, we select the brightest LAEs with a number density cut of

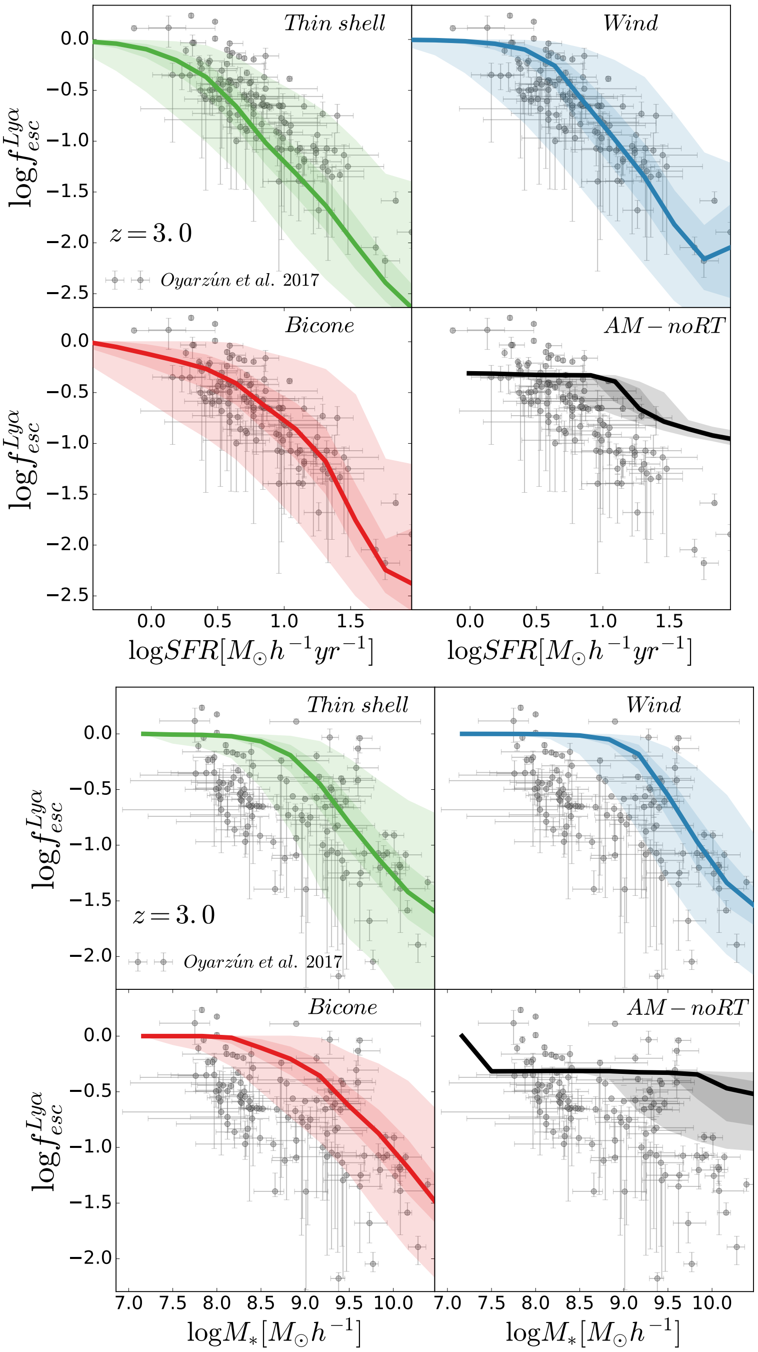


Fig. 2.7 The Lyα $f_{\rm esc}^{\rm Ly\alpha}$ as a function of SFR (top panels) and stellar mass (bottom panels) at $z = 3$. Gray points are from Oyarzún et al. (2017). Each panel displays our model predictions with a different outflow geometry, as shown in the legend. The bottom-right corner displays the predictions of the model with no radiative transfer. The solid line in each panel is the median of $f_{\rm esc}^{\rm Ly\alpha}$ predicted by our models. The dark and light coloured shaded regions display the $32 - 68$ and $5 - 95$ percentiles of the models predictions, respectively.

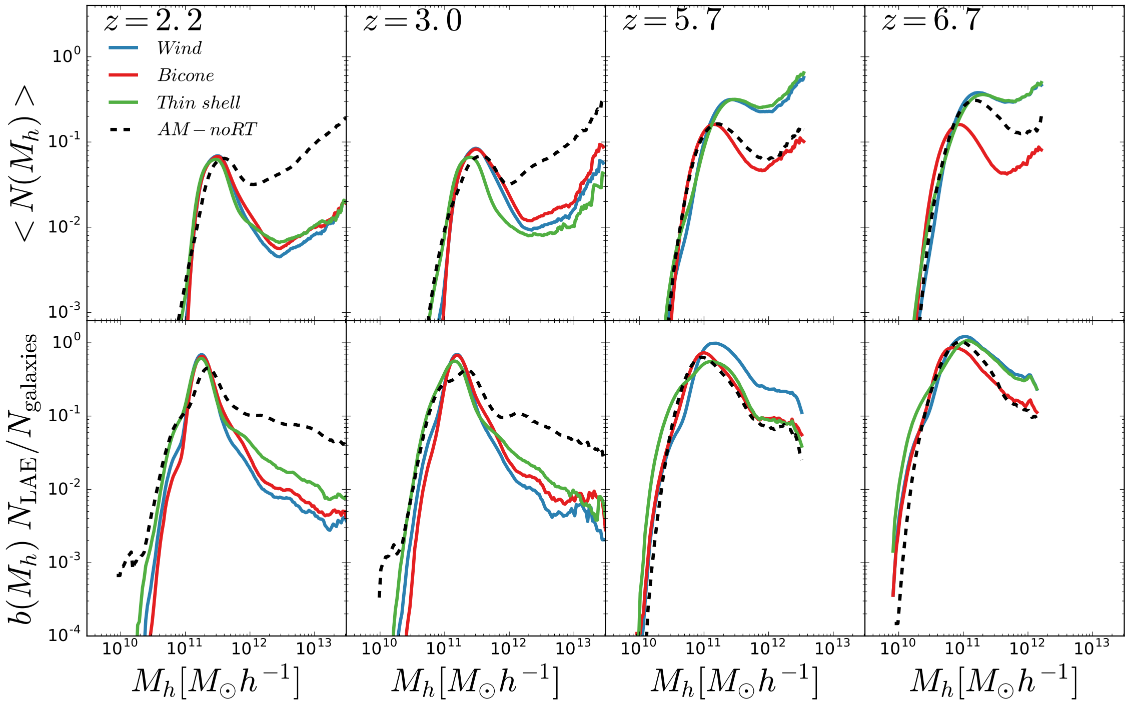


Fig. 2.8 **Top**: the halo occupation distribution (HOD) at redshift 2.2 , 3.0 , 5.7 and 6.7 from left to right. Model with radiative transfer show as blue, red and green solid lines for the *Wind*, *Bicone* and *Thin Shell* geometry respectively. The LAE sample AM-noRT is plotted as dashed black line. **Bottom**: fraction of galaxies that are considered LAE times the bias of the hosting dark matter halo. This quantifies the contribution of the different $\mathrm{M_h}$ to the overall bias of the population.

$10^{-3}\mathrm{h^3cMpc^{-3}}$.

Fig. 2.8 shows the halo occupation distribution (HOD) at $z = 2.2, 3.0, 5.7$ and 6.7. This is constructed by computing the mean number of galaxies within different halo mass bins. All models including radiative transfer display a similar HOD at $z = 2.2$ and 3.0. Central galaxies have a peak abundance in haloes of mass $\mathrm{M_{halo}} \approx 2 \times 10^{11}\mathrm{M_\odot h^{-1}}$. Satellite galaxies start dominating the abundance of haloes of mass $\mathrm{M_{halo}} \gtrsim 10^{12}\mathrm{M_\odot h^{-1}}$. None of the HODs at these redshifts reach $N(M_h) = 1$. Even at the peak of occupation, less than 10% of haloes host a LAE, regardless of radiative transfer effects.

At $z \geq 5.7$ the HOD of the *Bicone* model falls significantly below that from the *Thin Shell* and *Wind* models. This reflects the differences in the LFs at these high redshifts. As the *Bicone* model is not able to reproduce the observed LF the resulting

LAE population have quite different properties to the other RT samples.

The model with no radiative transfer systematically places LAEs in higher mass haloes compared to the radiative transfer models at low redshift. The occupation peak for centrals in the AM-noRT model is shifted to slightly more massive halos at $z = 2.2$ and 3.0. Additionally, at these redshifts, the occupation of dark matter halos with $\mathrm{M_h} \geq 10^{12}\mathrm{M_\odot h^{-1}}$ is much greater in the AM-noRT model than in the models including RT. At redshifts $z = 5.7$ and 6.7, the trend is inverted as LAE (*Thin Shell* and *Wind* geometry) populate halos slightly more massive than the AM-noRT model. Also, the occupation of halos with $\mathrm{M_h} \geq 10^{12}\mathrm{M_\odot h^{-1}}$ is greater in the RT models.

The bottom panels of Fig. 2.8 show the quantity

$$b(M_h)N_{\mathrm{LAE}}(M_h)/N_{\mathrm{galaxies}}(M_h), \tag{2.10}$$

where $N_{\mathrm{LAE}}(M_h)$ is the number of sources in our LAEs samples in a halo mass bin, $N_{\mathrm{galaxies}}(M_h)$ is the number of galaxies in the same M_h bin and the galaxy bias $b(M_h)$ is defined as

$$\xi_{\mathrm{galaxy}} = b^2\ \xi_{\mathrm{dark\ matter}}, \tag{2.11}$$

where ξ_{galaxy} and $\xi_{\mathrm{dark\ matter}}$ are the two point correlation functions for the galaxies and dark matter. This exhibits the contribution of different mass bins to the overall clustering bias of the LAE population. There is an evolution in the $\mathrm{M_h}$ that contributes to the bias, being greater at lower redshifts and lower at higher redshift. In particular, the peak values varies from $\mathrm{M_{halo}} \approx 2 \times 10^{11}\mathrm{M_\odot h^{-1}}$ at $z = 2.2$ to $\approx 6 \times 10^{10}\mathrm{M_\odot h^{-1}}$ at $z = 6.7$.

At low redshift ($z = 2.2$ and 3.0) the greater contribution to the bias come from lower mass halos in the RT models than in the AM-noRT model. However, this trend is inverted at z=5.7. Additionally, at $z = 6.7$ the main contribution to the bias comes from the same halo mass for all the models.

Overall there is a good agreement among some theoretical (e.g. Jose et al., 2013; Garel et al., 2015) and observational (e.g. Hagen et al., 2014; Kusakabe et al., 2018) work that predict/observed that LAE reside in dark matter halos with masses about $10^{11}\mathrm{M_\odot}$ and present stellar masses around $10^8 - 10^9\mathrm{M_\odot}$.

Table 2.3 Mock catalog characteristics including the redshift z, the redshift width Δz, sky coverage (Area), the size along the line of size $\mathrm{L}_{\parallel}$, the distance perpendicular to the line of sight $\mathrm{L}_{\perp}$, the number of mocks sliced from the simulation box $\mathrm{N}_{\mathrm{mocks}}$ and the median number of LAEs the mocks $\langle \mathrm{N}_{\mathrm{LAE}} \rangle$ with the 32 and 68 percentiles.

Authors	z	Δz	Area (deg^2)	$\mathrm{L}_{\parallel}$ (cMpc)	$\mathrm{L}_{\perp}$ (cMpc)	$\mathrm{N}_{\mathrm{mocks}}$	$\langle \mathrm{N}_{\mathrm{LAE}} \rangle$				
							Survey	Thin shell	Wind	Bicone	AM-noRT
Kusakabe et al. (2018)	2.2	0.0773	0.93	104.9	93.6	448	1248	1196^{+94}_{-90}	1191^{+95}_{-80}	1189^{+105}_{-78}	1183^{+109}_{-91}
Bielby et al. (2016)	3.0	0.0633	1.07	60.0	119.1	468	643	639^{+48}_{-57}	639^{+53}_{-51}	637^{+52}_{-60}	631^{+66}_{-59}
Ouchi et al. (2018)	5.7	0.0954	7.67	43.5	401.5	18	734	725^{+15}_{-9}	731^{+11}_{-19}	720^{+19}_{-21}	719^{+16}_{-20}
Ouchi et al. (2018)	6.7	0.1078	21.2	41.0	696.5	19	873	873^{+6}_{-30}	865^{+17}_{-19}	864^{+21}_{-20}	866^{+24}_{-6}

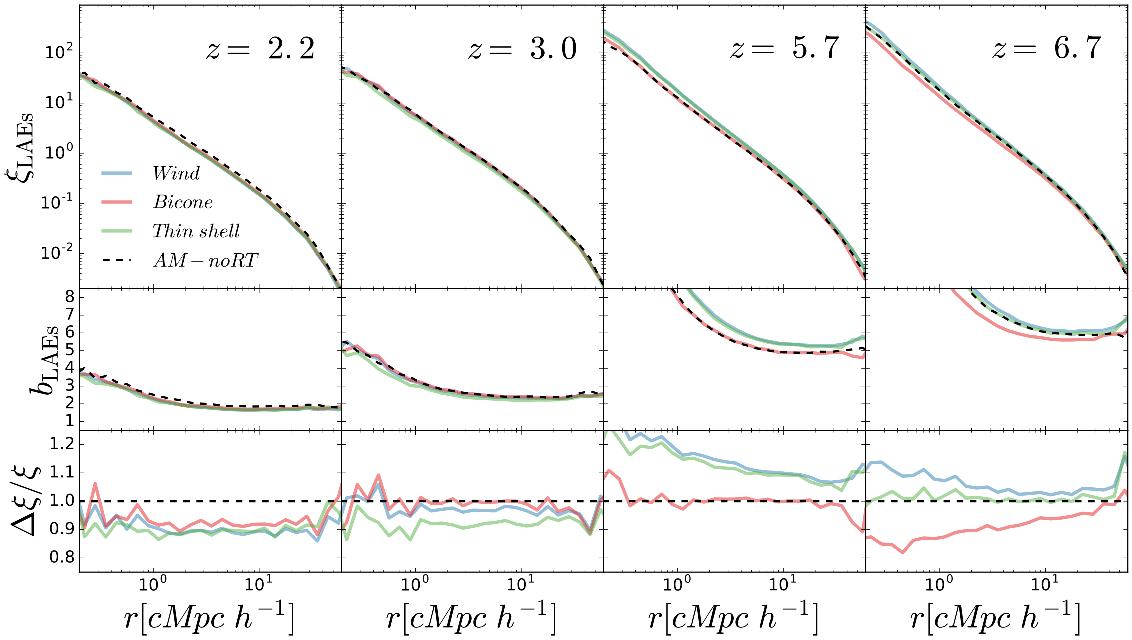


Fig. 2.9 **Top panels** : 3D auto-correlation function for the AM-noRT sample (black), for the *Thin Shell* (green), galactic wind (blue) and biconical galactic wind (red) for redshift 2.2 , 3.0 , 5.7 and 6.7 from left to right. **Middle panels** : The ratio between the different LAE sample and the dark matter correlation function. **Bottom panels** : relative difference between the different samples and the AM-noRT 3D correlation function.

2.4.6 The clustering of LAEs.

In this section we study how Lyα radiative transfer impacts the clustering of LAEs for each of the outflow geometries implemented. The sample used in this section is the same as the one used in §2.4.5.

In Fig. 2.9 the top panel shows the spherically-averaged 2-point auto-correlation function (2PCF) in real space at $z = 2.2, 3.0, 5.7$ and 6.7. The middle panel shows the bias, defined as in Eq.2.10. Moreover, in order to highlight the differences in the RT samples and the AM-noRT we show in the bottom panel of Fig.2.9 the relative difference of the 2PCF of the LAE samples $\xi_{\rm LAE}$ and the AM-noRT, i.e., $\Delta\xi/\xi = (\xi_{\rm LAE} - \xi_{\rm AM-noRT})/\xi_{\rm AM-noRT}$, where $\xi_{\rm AM-noRT}$ is the AM-noRT 2PCF.

Overall, the clustering amplitude increases towards higher redshifts regardless of the LAE model variant. In detail, each model predicts a slightly different clustering bias. There is a strong scale-dependence of the clustering bias in all models and at all redshifts for separations below $r \lesssim 15-20[\mathrm{Mpc}/h]$.

At $z = 2.2$ and 3.0 the clustering amplitude of the AM-noRT sample is about 10% above the one predicted by the RT models. This is a consequence of LAEs being hosted by higher mass dark matter halos for this model, as shown in previous sections. At $z = 5.7$ and 6.7, the clustering amplitude of the *Thin Shell* and *Wind* LAE samples are above that of the AM-noRT and *Bicone* models. Interestingly, as shown in the bottom panels of Fig. 2.9, towards redshifts $z > 3$ the AM-noRT sample features a slightly different slope with respect to the RT models.

In summary, the predicted clustering of LAEs at $z \lesssim 3$ is overall slightly lower when radiative transfer is included, and slightly higher towards $z \gtrsim 3$. The relative differences in the amplitude of clustering, with respect to the AM-noRT model, are of the order of 10%. These differences result from the non-trivial relation between the Lyα luminosity of galaxies and the dark matter halo population hosting these objects.

2.4.7 The clustering in mock catalogs of LAE surveys

In this section we compare our clustering prediction against several measurements of the clustering of LAEs at different redshifts from Kusakabe et al. (2018) at $z = 2.2$, Bielby et al. (2016) at $z = 3.0$ and Ouchi et al. (2010, 2018) at $z = 5.7$ and 6.7, respectively. We build LAE mock catalogs mimicking the properties of the different surveys to allow a close comparison with the observational datasets. These surveys use narrow band photometry to detect LAEs over a restricted redshift range. The main difference in the mock catalogs comes from the specific area, flux depth and equivalent width limit (EW) of the individual survey.

To build the mock catalogues, we choose a direction as line of sight (LoS). Assuming a distant observer, a galaxy coordinate is transformed in redshift space using

$$s = x_{\mathrm{LoS}} + \frac{\mathrm{v}_{\mathrm{LoS}}}{a(z)H(z)}, \tag{2.12}$$

where x_{LoS} is the galaxy coordinate along the LoS, $\mathrm{v}_{\mathrm{LoS}}$ is the galaxy peculiar velocity along the LoS and $a(z)$ and $H(z)$ are the scale factor and the Hubble parameter, respectively, at the Lyα pivot redshift, z_{pivot}, of the NB filter. Additionally, we conserve

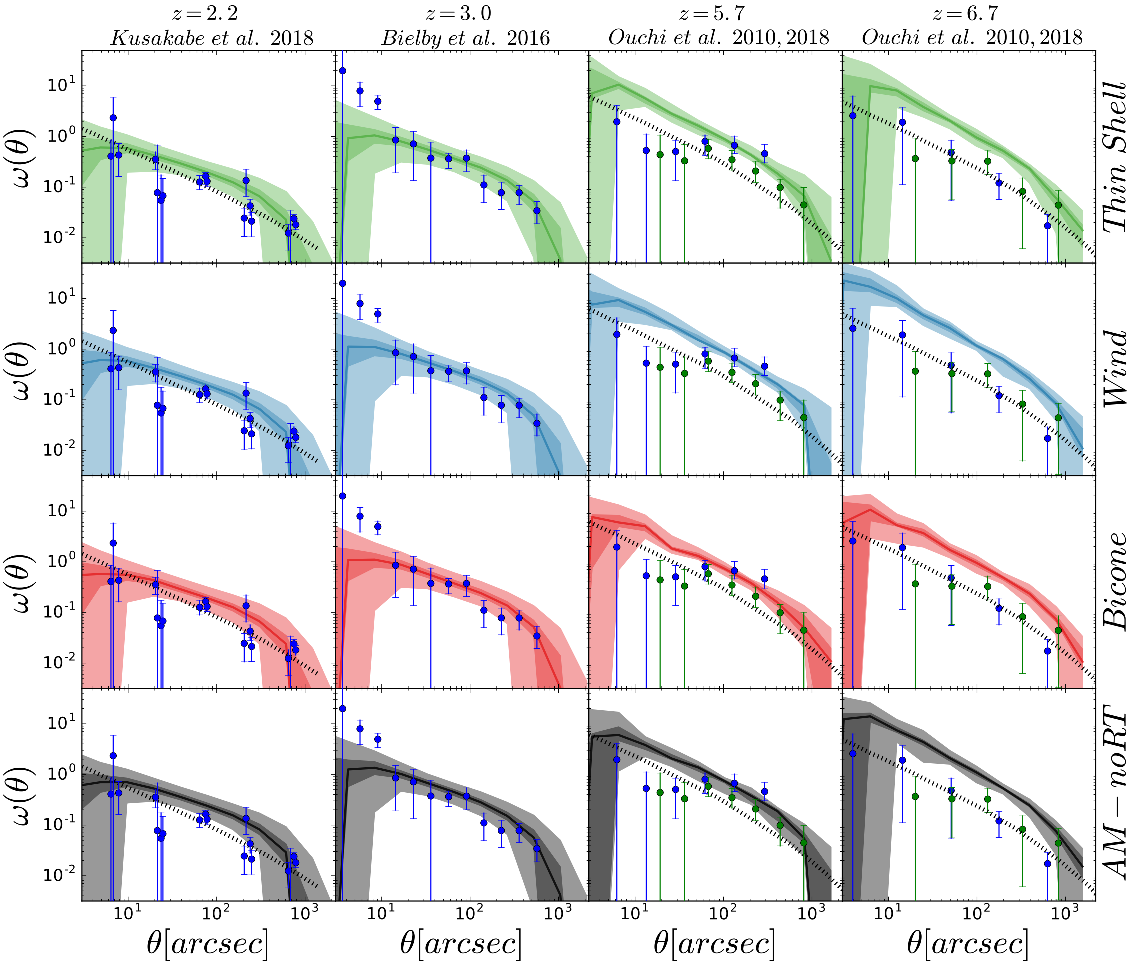


Fig. 2.10 Comparison between different model mocks (*Thin Shell*, *Wind*, *Bicone* and AM-noRT in rows from top to bottom) and the observed 2-point projected correlation function (Kusakabe et al., 2018; Bielby et al., 2016; Ouchi et al., 2010, 2018) at redshifts 2.2, 3.0, 5.7 and 6.7 in each column from left to right. The observational data is shown by dots and the best fitting power law $\omega(\theta)$ extracted from their original work are plotted as dashed black lines. The solid lines correspond to the median $\omega(\theta)$ for the mocks and the darker and lighter shades to the 32-68 and 5-95 percentiles respectively.

the periodicity of the box along the LoS direction.

Although some surveys have complicated footprints due to multiple pointings, our mocks are constructed as squares comprising an area equal to that of the target survey. Thus, the simulation box is simply split in slices along the LoS. The size perpendicular to the LoS is computed as

$$L_{\perp} = \sqrt{\mathrm{A}_{\mathrm{survey}}}, \tag{2.13}$$

where $\mathrm{A}_{\mathrm{survey}}$ is the survey sky coverage. The thickness (along the LoS) of the slice is computed as

$$L_{\|} = D_{\mathrm{co}}(z = z_{+}) - D_{\mathrm{co}}(z = z_{-}), \tag{2.14}$$

where $D_{\mathrm{co}}(z)$ is the comoving distance at the geometric redshift z. Additionally,

$$z_{\pm} = \frac{\lambda_p \pm 0.5\ \mathrm{FWHM}}{\lambda_{Ly\alpha}} - 1, \tag{2.15}$$

where λ_p and FWHM are the pivot wavelength and the full width half maximum of the narrow band filter and $\lambda_{Ly\alpha}$ is the Lyman α wavelength.

We calculate the limiting luminosity L_{cut} and the minimum rest frame equivalent width $\mathrm{EW}_{0,\mathrm{cut}}$ for each survey by matching the LAE number density, n_{LAE} of the surveys to the one in the whole simulation box (see Appendix 2.7.3). Then, our mock catalogs consist of galaxies with luminosity above L_{cut} and EW_0 above $\mathrm{EW}_{0,\mathrm{cut}}$. Table 2.3 lists the properties of the mocks, including the parallel and transverse sizes along the LoS, the redshift window $\Delta z = z_{+} - z_{-}$, the number of mocks, $\mathrm{N}_{\mathrm{mock}}$, sliced from the simulation box and the number of LAE in each survey, and the median with 32-68 percentiles of the number of LAEs in the mocks.

The value of $L_{\|}$ for narrow-band surveys is typically very small compared to the box length of the simulation. This allows for a big fragmentation of the simulation box along the LoS. On the other hand, $L_{\perp}$ can vary significantly between surveys. While, at low redshift ($z = 2.2, 3.0$) $L_{\perp}$ is relatively small and allows a large number of mock surveys, at $z = 5.7, 6.7$ only one cut is possible due to the large size required for the mock surveys. As a result of this, the number of mocks at $z = 2.2, 3.0$ (448 and 468 respectively) is much larger than that at $z = 5.7, 6.7$ (18 and 19 respectively).

Since n_{LAE} in the simulation box is set to match the observed n_{LAE} of each survey (see Appendix 2.7.3), the observed number of LAE and the median number of LAE in our mocks, $\langle N_{\mathrm{LAE}}\rangle$, are compatible within 1 sigma. Additionally, the dispersion of $\langle N_{\mathrm{LAE}}\rangle$ is higher (lower) at $z = 2.2$ and 3.0 (5.7 and 6.7), since the comoving volume is smaller (larger). Hence, the impact of cosmic variance on clustering measurements is stronger (weaker).

We construct mock catalogs of LAE surveys from Kusakabe et al. (2018) at $z \approx 2.2$, Bielby et al. (2016) at $z \approx 3$, Ouchi et al. (2010) at $z \approx 5.7$ and Ouchi et al. (2018) at $z \approx 6.7$. Figure 4.20 shows the comparison between the observed angular 2-point correlation function of these surveys, $\omega_{\rm survey}$, and that computed from the mock catalogues, $\omega_{\rm mock}$. Note that in Bielby et al. (2016) the authors rescaled their observed clustering by a factor of 1.64 to correct for contaminants. This kind of correction was not made in the other works used in this section. Here we use the Bielby et al. (2016) data set including their correction.

Overall, $\omega_{\rm mock}$ is very similar among our different model variations, including the AM-noRT model. The differences in the clustering due to the different bias of the samples are small in comparison with the scatter due to cosmic variance, making all models indistinguishable from each other.

At redshift 2.2 and 3.0 there is a good agreement between the the mocks and the clustering measurements in Kusakabe et al. (2018) and Bielby et al. (2016) respectively. At higher redshifts the LAE clustering predicted by the mocks is overestimated in our models. In particular, at $z = 5.7$, for angular distances $\theta < 50$ arcsec, $\omega_{\rm mock}$ overestimates the clustering, while at larger θ the mocks match very well $\omega_{\rm survey}$. Additionally, at redshift 6.7 the $\omega_{\rm mock}$ bias is significantly (about 2-sigma) overestimated in comparison with $\omega_{\rm survey}$. This discrepancy could be caused by multiple reasons. The moderate contamination of interlopers ($\sim 10\%$) in the Ouchi et al. (2010) sample could decrease the measured clustering amplitude. Also, the observed LAE population at this redshift might contain a significant contribution of objects at the mass resolution limit of our `GALFORM` galaxies ($M_{\rm halo,min} \approx 10^{10}[{\rm M}_{\odot}/{\rm h}]$, see §2.2.3), thus making our predictions biased towards higher masses and clustering amplitudes. In fact, we have checked that about a 10% of the LAE in our different samples have halo masses below $10^{10.5}[M_{\odot}/h]$.

2.5 Discussion.

Here we discuss some of the results found in previous sections. In particular, in subsection 2.5.1 we discuss how the different outflow geometries impact the predicted properties of the LAE populations. Then, in subsections 2.5.2 and 2.5.3 we discuss the limitations of our methodology.

Table 2.4 Fraction of shared galaxies between pairs of models at the same redshift.

z	Model	Thin Shell	Wind	Bicone	AM-noRT
2.2	Thin Shell	1.000	0.814	0.555	0.229
	Wind	0.814	1.000	0.592	0.189
	Bicone	0.555	0.592	1.000	0.197
	AM-noRT	0.229	0.189	0.197	1.000
3.0	Thin Shell	1.000	0.805	0.401	0.188
	Wind	0.805	1.000	0.427	0.151
	Bicone	0.401	0.427	1.000	0.227
	AM-noRT	0.188	0.151	0.227	1.000
5.7	Thin Shell	1.000	0.413	0.798	0.108
	Wind	0.413	1.000	0.322	0.063
	Bicone	0.798	0.322	1.000	0.160
	AM-noRT	0.108	0.063	0.160	1.000
6.7	Thin Shell	1.000	0.354	0.663	0.104
	Wind	0.354	1.000	0.229	0.076
	Bicone	0.663	0.229	1.000	0.259
	AM-noRT	0.104	0.076	0.259	1.000

2.5.1 Differences between the RT models.

In this work we have used three different gas outflow geometries (*Thin Shell*, spherical galactic wind and biconical galactic wind) to model the Lyα radiative transfer inside galaxies. The galaxy properties predicted for LAEs are very similar. The only significant difference between the predictions of different geometries is on the required distributions of column density and expansion velocity.

In Table 2.4 we list the fraction of galaxies shared by pairs of LAE models imposing $\mathrm{EW}_0 > 20$Å and a number density cut of $10^{-3}\mathrm{h}^3\mathrm{cMpc}^{-3}$. We find that the *Wind* and *Thin Shell* geometries share a high fraction of galaxies ($\sim 80\%$) at redshifts 2.2 and 3.0. However, at high redshift these geometries select different galaxies as the shared fraction in relatively low ($\sim 40\%$ overlap). This might be due to the fact that there is a necessity of $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha} \sim 1$ and the recipes to compute $\mathrm{N}_{\mathrm{H_I}}$ and $\mathrm{V}_{\mathrm{exp}}$ are different. However, quite the opposite relation is seen between the *Thin Shell* and *Bicone*, as at low redshift they share a relatively low percentage of galaxies ($\sim 45\%$) and this increase at higher redshifts ($\sim 70\%$).

Finally, when comparing the galaxies in the *Wind* and *Bicone* geometry we surprisingly find a low overlap between them. In particular, the maximum overlap happens at

$z = 2.2$ ($\sim 55\%$) and it drops down to only $\sim 20\%$ at $z = 6.7$. This shows the impact of the gas geometry on how the RT shapes the LAE selection function; even though the intrinsic galaxy population and the recipes to derive $\mathrm{N_H}$ and $\mathrm{V_{exp}}$ are the same, the two geometries predicts different populations (although with similar characteristics).

We conclude that the RT LAE samples, in general, share a big fraction of galaxies ($\geq 50\%$) although the implemented gas geometries are very different. This is due to the fact that $f_{\rm esc}^{\rm Ly\alpha}$ behaves similarly for all of them. In particular, even if the exact dependence is different for each geometry, decreasing $\mathrm{N_H}$, increasing $\mathrm{V_{exp}}$ and decreasing τ_a increase $f_{\rm esc}^{\rm Ly\alpha}$ thus the visibility of the object for all of them. This makes the RT LAE samples very similar, as galaxies with properties that maximize $\mathrm{L_{Ly\alpha}}$ and $f_{\rm esc}^{\rm Ly\alpha}$ are selected.

2.5.2 Limitations of the simple AM-noRT model.

We have also used a very simplistic LAE model where radiative transfer effects are not taken into account and $\mathrm{L_{Ly\alpha}}$ depends monotonically on the SFR. In Table 2.4 we also list the overlap between the radiative transfer and AM-noRT LAE sample. We find that the fraction of galaxies shared between the AM-noRT and RT catalogs is low, reaching its maximum value at z=2.2 ($\sim 20\%$) and then decreasing to $\sim 7\%$ at redshift 6.7.

As shown in Fig. 2.7 the AM-noRT sample fails to match not only the observed $f_{\rm esc}^{\rm Ly\alpha}$–SFR and $f_{\rm esc}^{\rm Ly\alpha}$–$M_*$ relations but also the overall trend where $f_{\rm esc}^{\rm Ly\alpha}$anti-correlates with these two properties due to the RT (as described above). Additionally, the dark matter halo population, and thus the clustering, is different in comparison with the RT samples.

This work highlights the importance of taking into account the Lyα RT inside galaxies when modeling LAEs. In particular, unlike in RT LAE samples, the galaxy properties model AM-noRT differ from observations, making them less attractive to study galaxy formation and evolution.

2.5.3 Limitations of the RT models.

The IGM impact on Lyα detection in still unclear. However, some works have explored that the IGM plays a major role in the detectability of galaxies based on Lyα flux (Dijkstra et al., 2006; Zheng et al., 2011; Behrens et al., 2017). The IGM opacity becomes more important at higher redshifts (~ 7) where the universe is denser and colder. However, the IGM might already also have an impact on the LAE selection function at $z = 2.2$ as, even if the universe is highly ionized, the cross-section of neutral hydrogen atoms for scattering Lyα photons is very high. The IGM impact might alleviate some of the tension that we find when we compare LAE models with observations. We will implement the effect of the IGM opacity in future work.

In Fig. 2.7 we found that, although the observed $f_{\rm esc}^{\rm Ly\alpha}$–SFR relation is perfectly reproduced by our RT models, the $f_{\rm esc}^{\rm Ly\alpha}$–M$_*$ relation is not. Even if the overall trend is similar, we find a significant difference (about 0.5 dex) in the stellar mass. This is probably not caused by our implementation of RT in a semi-analytic model, but by `GALFORM` itself, as we note that full `GALFORM` $\rm M_h - M_*$ relation at redshift 3.0 is overestimated (also about 0.5 dex) in comparison with the observed one (Behroozi et al., 2010). Another possible source for this discrepancy is the different stellar population synthesis models used by Oyarzún et al. (2017) and `GALFORM`.

Another limitation of the RT models is that they predict very similar galaxy properties for the three different geometries. This degeneracy makes it difficult to determine from observations which geometry is the one driving the Lyα photons escape. Nonetheless, the three gas geometries used in this work have very different Lyα line profiles (as shown in figure 2.2) which might break the degeneracies and lead to a better understanding of the escape channels of Lyα radiation. We will implement line profiles in a upcoming work.

2.6 Conclusions and future work.

Lyman-α emitters are a promising galaxy population to trace the large scale structure of the Universe at high redshifts, $z \gtrsim 2$. One of the main advantages of LAEs is their high luminosity at the Lyα rest frame wavelength, making them easy to detect. Additionally, due to the Hubble expansion (Hubble, 1929), the Lyα line is observable in the optical from $z \sim 2$ to ~ 7, allowing ground-based measurement of these galaxies.

However, their selection function is quite complex as it depends upon Lyα radiative transfer, which is sensitive to local astrophysical conditions.

We have designed a theoretical model of LAE based on a Monte Carlo Radiative Transfer code that can be applied to huge cosmological volumes. In particular, we have applied our model the N-body only-dark-matter simulation `P-Millennium` and the semi-analytical model of galaxy formation and evolution `GALFORM` (Lacey et al., 2016).

Monte Carlo Radiative Transfer codes have demonstrated to be a powerful tool to understand how Lyα photons escape from galaxies. Unfortunately, the high computational cost prohibits the capability of being directly run over cosmological volumes. In order to avoid this problem we have developed analytical expressions for the Lyα escape fraction $f_{\rm esc}^{\rm Ly\alpha}$ that are quite accurate for a wide range of outflow expansion velocities $V_{\rm exp}$, neutral hydrogen column densities N_H and metallicities Z.

Our methodology computes $f_{\rm esc}^{\rm Ly\alpha}$ for each galaxy as a function of Z, $V_{\rm exp}$ and $N_{\rm H_I}$, which characterise the gas outflows from which Lyα photons escape. We compute these quantities using galaxy properties such as the size, SFR or halo mass. Free parameters to compute these quantities are chosen to fit the observed luminosity function over a wide range of redshifts. After calibration we find that every geometry reproduces well the observed LAEs LF at low redshift while only the *Thin Shell* and *Wind* manage to match them at high redshift. We conclude that our *Bicone* geometry (as described in this work), at high redshift, is less favored with respect to the others. In a future work we will redesign this geometry by making it more realistic and including a dependence in the $f_{\rm esc}^{\rm Ly\alpha}$ and other properties with the line of sight.

We have analyzed the relative abundance of Lyα emitters by breaking down their LF in terms of several properties. Halo or stellar masses are not significantly correlated with Lyα luminosities. The LF is actually mostly dominated by relatively low mass galaxies. However, when the LF is split in SFR bins we find a clear positive correlation with Lyα luminosity. Finally, when the LF is divided into metallicity bins we find a scattered correlation for $\log(Z) < -2$. Moreover, the contribution of high metallicities ($\log(Z) > -2$) to the bright end of the LF is small.

We also compared the properties of a Lyα selected sample to the bulk of the galaxy population at high redshifts. We find that LAEs lie in relatively low mass halos. Additionally, the galaxies with the strongest starburst episodes are not selected as LAE since these galaxies typically have higher metallicities, and thus their $f_{\rm esc}^{\rm Ly\alpha}$ is low.

To validate our predicted $f_{\rm esc}^{\rm Ly\alpha}$, we have compared our LAE samples to the observational data from Oyarzún et al. (2017). We find a remarkable good agreement between our predictions and the observationally measured $f_{\rm esc}^{\rm Ly\alpha}$ - SFR relation. The LAE samples including RT reproduce successfully this anti-correlation and the scatter found between these quantities. However, the predicted $f_{\rm esc}^{\rm Ly\alpha}$ - M_* plane is offset by ~ 0.5 dex in M_* with respect to the data from Oyarzún et al. (2017). This difference can be due to the different assumptions about the stellar population synthesis models used by Oyarzún et al. (2017) and `GALFORM`, the impact of a different IMF in `GALFORM`, or simply that `GALFORM` predicts significantly more massive star-forming galaxies at these higher redshifts with respect to observational estimates.

Finally, we find that our LAE AM-noRT sample based on assuming a monotonic relation between SFR and $\rm L_{Ly\alpha}$ is not able to reproduce any of the observed trends. This highlights the crucial role of RT in shaping the LAE selection function.

We have also studied the dark matter halo population hosting LAEs in our models. We find differences between the samples including RT and the sample without RT. At low redshift, in comparison with the AM-noRT, the RT models predicts lower mass dark matter halos host LAE. This trend reverses at high redshift, as LAEs lie in more massive halos in the RT samples. We also find that the satellite fraction is low at all redshifts ($\sim 2\%$) and similar for all of the model variants.

The difference in the DM halo populations is directly translated into clustering discrepancies between the AM-noRT and RT samples. At low redshift, as a consequence of LAEs modeled with RT lying in lower mass DM halos, we find that they have a lower galaxy bias than the AM-noRT sample. This trend is reversed at high redshifts, when RT LAEs lie in more massive dark matter halos. Thus, we find that the RT models have a steeper galaxy bias evolution than the model excluding RT.

Finally, we have compared our model clustering predictions with observations finding some tension. While at redshifts 2.2 and 5.7 the observed clustering is well reproduced, at redshifts 3.0 and 6.7 the galaxy bias is poorly constrained. Although the impact of the intergalactic medium on Lyα is not very clear, previous works (Zheng et al., 2011) indicate that the IGM transmission could have an impact on Lyα selected samples that might alleviate this tension.

We have demonstrated the importance of RT in shaping the selection function of LAEs for galaxy properties as metallicity, SFR or DM halo properties. On one hand, the peculiar observational trends found can not be reproduce with a simple monotonic relation between SFR and $L_{Ly\alpha}$. On the other hand, the inclusion of RT changes in a very particular way the clustering of Lyα selected samples. All this make extremely important to construct models with Lyα RT in order to understand the galaxy properties, formation and evolution of LAEs. Moreover, future surveys tracing the large scale structure of the Universe through LAEs will require a deep understanding of the channels through which Lyα photons escape in order to obtain unbiased cosmological constrains.

In future work we plan to implement the transmission of Lyα photons through the IGM, which is especially important at high redshifts. In order to do so we will develop analytic expression for the Lyα line profile and a model to compute the IGM transmission in large cosmological volumes. These tools will enable us to explore how the IGM shapes the LAE galaxy properties and clustering.

2.7 Extra material

2.7.1 Further details about the bicone geometry

In general, $f_{\rm esc}^{\rm Ly\alpha}$ can be define for an arbitrary angular aperture $\Delta\gamma$ as

$$f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma) = \frac{N_{\rm escaped}(\Delta\gamma)}{N_{\rm emitted}(\Delta\gamma)}, \tag{2.16}$$

where $N_{\rm emitted}(\Delta\gamma)$ is the number of photons emitted towards the aperture $\Delta\gamma$ and $N_{\rm escaped}(\Delta\gamma)$ is the number of photons that escape through that set of directions. Note that for the thin shell and the non-biconical galactic wind this is expression is valid and $\Delta\gamma$ is every direction (4π).

The bicone geometry is not spherical symmetry, thus its $f_{\rm esc}^{\rm Ly\alpha}$ depends on the line of sight. In the case of the bicone, the escape fraction presented in our work corresponds to the $f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma_b)$ where $\Delta\gamma_b$ is the angular aperture $\theta < \pi/4 \cup \theta > 3\pi/4$ and $\phi \in [0, 2\pi)$.

In other words, we define the biconical escape fraction as the number of photons that escape through the bicone divided by the number of photons emitted towards the bicone. We checked that the escape fraction within $\Delta\gamma_b$ does not change. Meanwhile, the escape fraction within $4\pi - \Delta\gamma_b$ is also constant and considered ~ 0.

In this formalism, the total escape fraction (taking into account ever possible line of sight) would be

$$f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma = 4\pi) = \frac{N_{\rm escaped}(\Delta\gamma_b) + N_{\rm escaped}(4\pi - \Delta\gamma_b)}{N_{\rm emitted}(\Delta\gamma = 4\pi)}, \tag{2.17}$$

where the aperture $\Delta\gamma = 4\pi - \Delta\gamma_b$ encapsulates every line of sight outside the bicone. Then, if we assume that no photon escapes outside the bicone, i.e, $N_{\rm escaped}(4\pi - \Delta\gamma_b) = 0$,

$$f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma = 4\pi) = \frac{N_{\rm escaped}(\Delta\gamma_b)}{N_{\rm emitted}(\Delta\gamma = 4\pi)}, \tag{2.18}$$

where if we assume that the photon emission is isotropic

$$f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma = 4\pi) = \frac{N_{\rm escaped}(\Delta\gamma_b)}{N_{\rm emitted}(\Delta\gamma_b)4\pi/\Delta\gamma_b} = f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma_b) \times \frac{\Delta\gamma_b}{4\pi}. \tag{2.19}$$

Note that, by construction, in our bicone implementation, $f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma = 4\pi) \leq \Delta\gamma_b/4\pi$, even when the dust content of the bicone is null. This is caused by the dusty optical thick disk in the plane perpendicular to the bicone axis. The disk absorbs every photon that enters into it. This includes i) photons sent directly by the source and ii) photons that are initially emitted towards the bicone, bound in its surface and end up in the disk. In this way, only when the bicone neutral hydrogen column density is 0, then $f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma = 4\pi) = \Delta\gamma_b/4\pi$.

To include the angular dependence of $f_{\rm esc}^{\rm Ly\alpha}$ we should i) assign a random orientation to every galaxy in our simulation and ii) compute the observed luminosity taking into account the orientation, i.e.,

$$L_{\rm Ly\alpha}(\Delta\gamma) = L^0_{\rm Ly\alpha} \times f_{esc}(\Delta\gamma), \tag{2.20}$$

where $\Delta\gamma = \Delta\gamma_b$ when the bicone is pointing towards the observer and $\Delta\gamma = 4\pi - \Delta\gamma_b$ when the line of sight does not go through the outflow.

By construction, a fraction of $1 - \Delta\gamma_b/4\pi$ galaxies would be seen through the optical thick disk and be assigned $f_{\rm esc}^{\rm Ly\alpha}(4\pi - \Delta\gamma_b) \sim 0$. This would decrease the number counts to LAEs too much in comparison with current observations. To counterbalance this we assigned $f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma_b) \sim 0$ to every galaxy, thus

$$L_{\rm Ly\alpha} = L^0_{Ly\alpha} \times f_{\rm esc}^{\rm Ly\alpha}(\Delta\gamma_b). \tag{2.21}$$

This imposes that the biconical outflow of every galaxy is pointing to the observer. Therefore our biconical model is an upper limit of the bicone geometry presented in this work. We are aware that this introduces some limitations in the conclusions of the bicone geometry, but the other two geometries are not affected by this. We plan to implement in a future work a improved biconical geometry and the angle dependence of its $f_{\rm esc}^{\rm Ly\alpha}$.

2.7.2 Validating the $f_{\rm esc}^{\rm Ly\alpha}$ fitting formulae

A direct comparison between the Monte Carlo radiative transfer (MCRT) code output and or $f_{\rm esc}^{\rm Ly\alpha}$ model for the *Thin Shell* geometry is shown in Fig. 2.12, for the galactic wind in Fig. 2.13 and for the *Bicone* in Fig. 2.14. These figures are divided in 8 panels sub divided in another two: a) escape fraction v.s. dust optical depth and b) the relative difference between our model and the radiative transfer code output. In panels a) the output of the radiative transfer code for a fixed N_H is plotted with $\rm V_{exp}$ color coded in solid lines with their respective errors (same colored shade region) computed using eq. 2.6 and our $f_{\rm esc}^{\rm Ly\alpha}$ model is plotted in black solid lines. In type b) panels we show the relative difference between our model and the MCRT code with the same color code than above.

In general, the performance of our model decrease with τ_a, this is because, as discussed above, decreasing $f_{\rm esc}^{\rm Ly\alpha}$ increases the errors. This disagreement, in some cases, leads to an overestimation of $f_{\rm esc}^{\rm Ly\alpha}$ when its true value is $\lesssim 0.01$. For our work, these low values are very rare and so do not affect our results. Overall, we find that the typical discrepancies are below 10% and 1% for $\log\tau_a < -0.5$ and < -1.0 respectively.

Our model for the *Thin Shell* $f_{\rm esc}^{\rm Ly\alpha}$ is able to reproduce the whole velocity range of our grid for $\rm N_H < 10^{19.5} cm^{-2}$, reaching a 99% accuracy in most cases.

Our $f_{\rm esc}^{\rm Ly\alpha}$ model using the *Wind* geometry is also able to reproduce the output of our radiative transfer code through most of the parameter, only failing at very high $\rm N_H$ and low V_{exp} combinations, where $f_{\rm esc}^{\rm Ly\alpha} < 0.1$. As in the *Thin Shell* geometry, our model behaves better for $\rm N_H > 10^{19.5} cm^{-2}$. In particular, in most of out grid, the disagreement is lower than 10% and for low $\tau_a < -1$ the typical agreement is 1%.

The *Bicone* geometry is more complex than the other geometries, and its $f_{\rm esc}^{\rm Ly\alpha}$ model has the worst performance of all. However, for most of the grid the model is within 10% errors. As explained in §2, the maximum $f_{\rm esc}^{\rm Ly\alpha}$ depends on the properties of the outflowing gas, causing that only systems with very low optical depth (low $\rm N_H$ and/or high $\rm V_{exp}$) manage to reach $f_{esc}^{Ly\alpha} = 1$. This also causes that in very optically thick systems $f_{\rm esc}^{\rm Ly\alpha}$ reaches 0.001 (even if there is no dust). We decided not to include $\rm N_H = 10^{22.5} cm^{-2}$ in our model because the maximum value of $f_{\rm esc}^{\rm Ly\alpha}$ at $\rm V_{exp} = 1000$ km/s would be about 0.01 and, as discussed above, it is unnecessary to reproduce such low values.

We show the distribution of dust optical depth τ_a against $\rm V_{exp}$ (top panels of Fig.2.11) and $\rm N_H$ (bottom panels of Fig.2.11) for our RT LAE samples (selected as in §2.4.1) at redshifts 2.2 , 3.0 , 5.7 and 6.7 from left to right. At redshifts 2.2 and 3.0 the *Wind* and *Thin Shell* τ_a distributions are very similar in width and center ($\log \tau_a \sim -2$) while the *Bicone* model predict $\log \tau_a \sim -3$. Since the *Bicone* $f_{\rm esc}^{\rm Ly\alpha}$ exhibits an upper limit < 1, it requires low column densities (see Fig. 2.4) and Z, thus low τ_a values. At high redshifts (5.7 and 6.7) the dust optical depth distributions for the three geometries are very similar and peak at $\log \tau \sim -3$. In addition to the bulk of the distribution, the *Thin Shell* and *Wind* geometries also present a small bump around $\log \tau \sim -0.5$. Finally, we have checked that the fraction of galaxies in our LAE samples is $< 2\%$ for all the configurations studied in this work.

In the top panels of Fig.2.11 we explore the evolution in redshift of the $\rm \tau_a - V_{exp}$ distributions of our LAE samples. While at low redshift ($z = 2.2$, 3.0) the three geometries occupy different regions in this plane, they merge at higher redshifts ($z = 5.7$, 6.7). At all redshifts the *Bicone* needs the lowest τ_a possible as the typical $f_{\rm esc}^{\rm Ly\alpha}$ in this geometry is very low. In the case of the *Thin Shell* and *Wind* the distribution passes

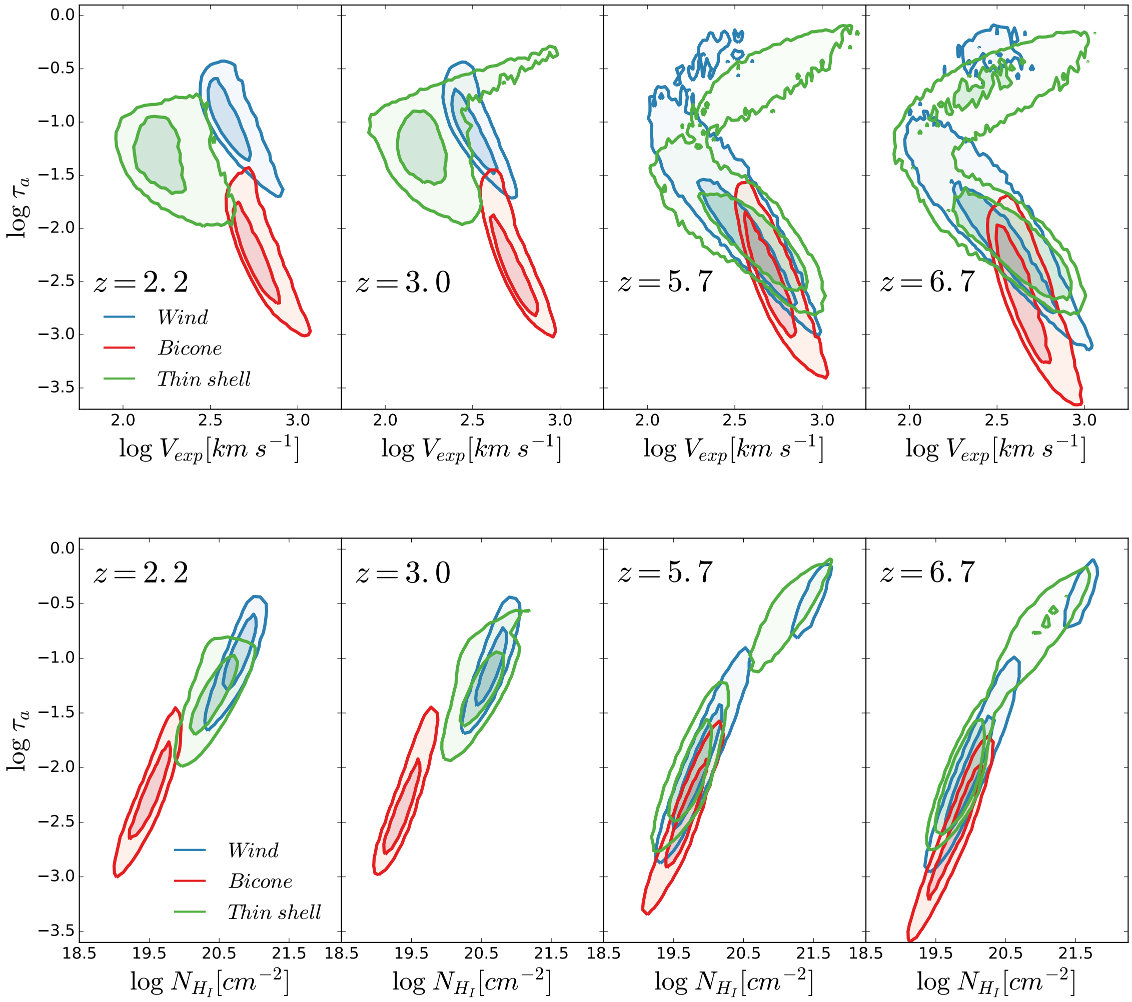


Fig. 2.11 Top : Same as Fig. 2.4 but displaying the dust optical depth τ_a and expansion velocity $V_{\rm exp}$ for the LAE samples with RT. Bottom: Same as Fig. 2.4 but displaying the dust optical depth τ_a and neutral hydrogen column density $N_{\rm H}$ for the LAE samples with RT.

from compact at low redshift to bimodal as also studied in Fig. 2.4 and explained in §2.4.1.

In the bottom panels of Fig.2.11 we explore the evolution in redshift of the $\tau_{\rm a} - N_{\rm H}$ distributions of our LAE samples. We see a clear correlation between these two quantities at all redshifts that is driven by Eq.4.3. Additionally, we see the dame behavior of $N_{\rm H}$ as in §2.4.1.

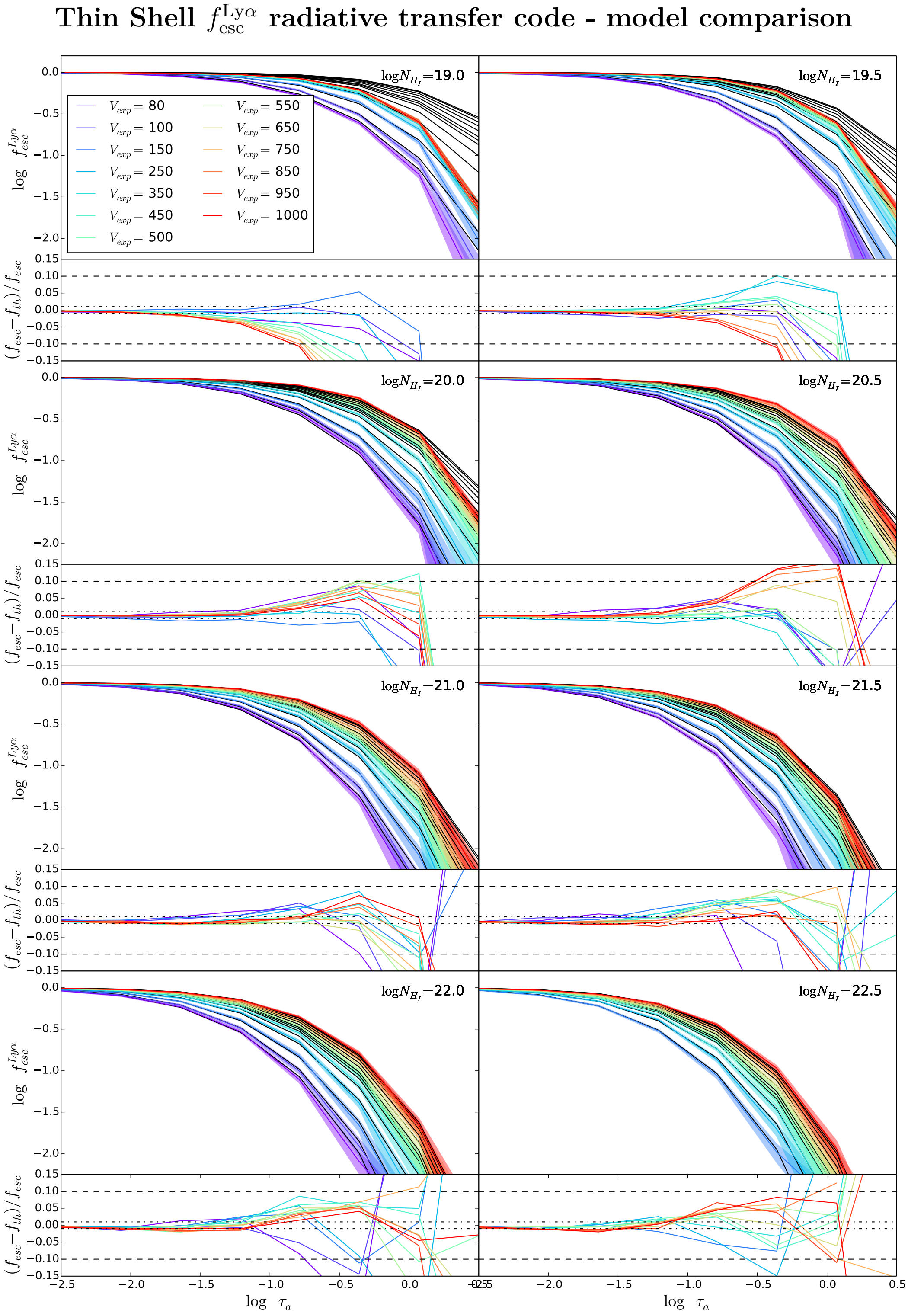
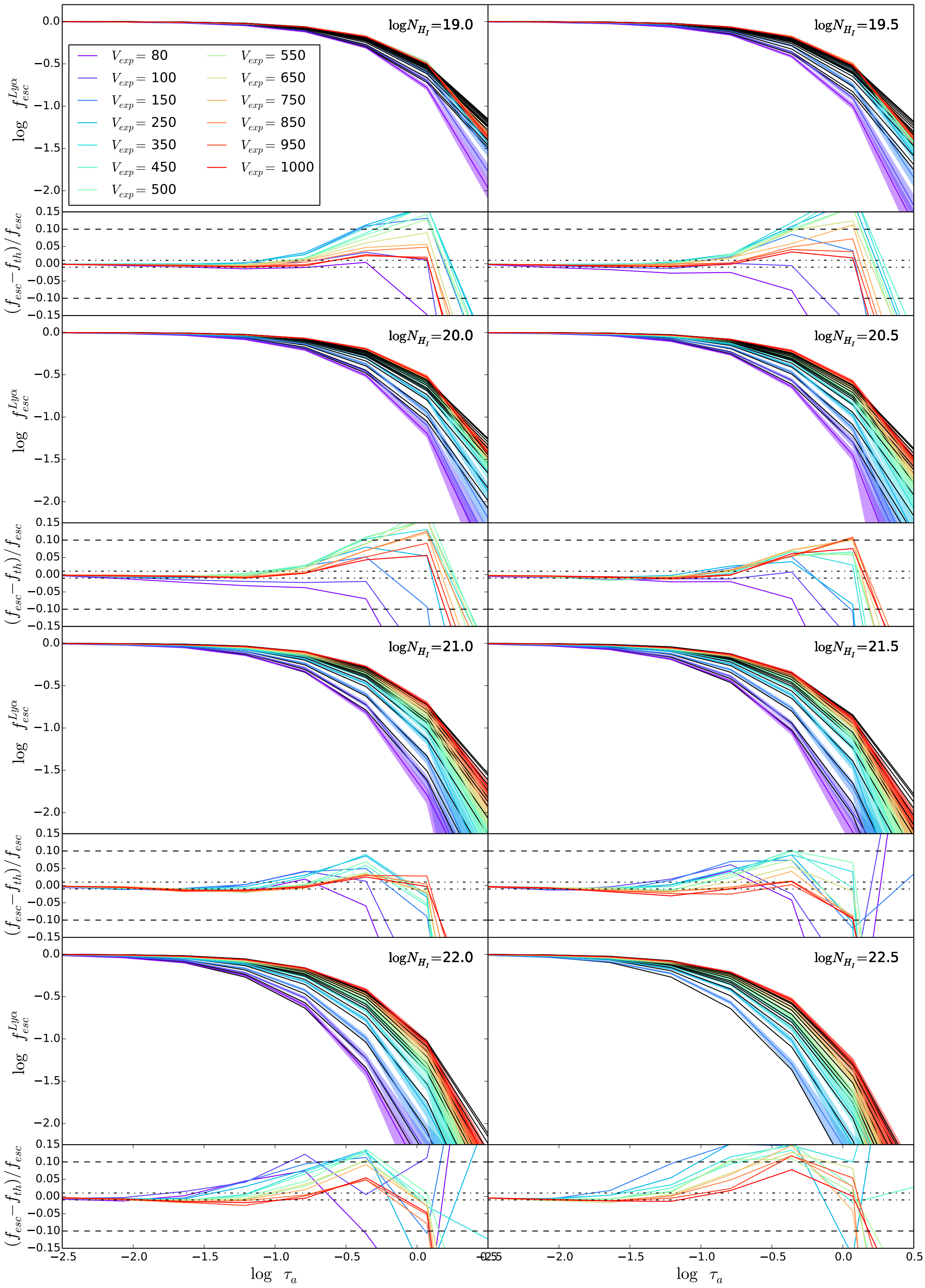

Thin Shell $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ radiative transfer code - model comparison

Fig. 2.12 Comparison between the output of the radiative transfer code and our model for the $f_{esc}^{Ly\alpha}$ in the *Thin Shell* geometry. Each panel is divided in top (the values of the escape fraction) and bottom (relative difference between our model and the radiative transfer code). In top panels the output from the radiative transfer code in plotted in colored lines (color coded by the velocity of the system) with their errors (shades with the same color) and our model prediction in black. In bottom panels the relative difference between our model and our code are plotted in colored lines and the ±1% and ±10% are represented by black dashed-dotted and dashed lines respectively. Note that he color code is the same in every panel.



Galactic wind $f_{\rm esc}^{\rm Ly\alpha}$ radiative transfer code - model comparison

Fig. 2.13 Same as figure 2.13 but for the galactic wind.

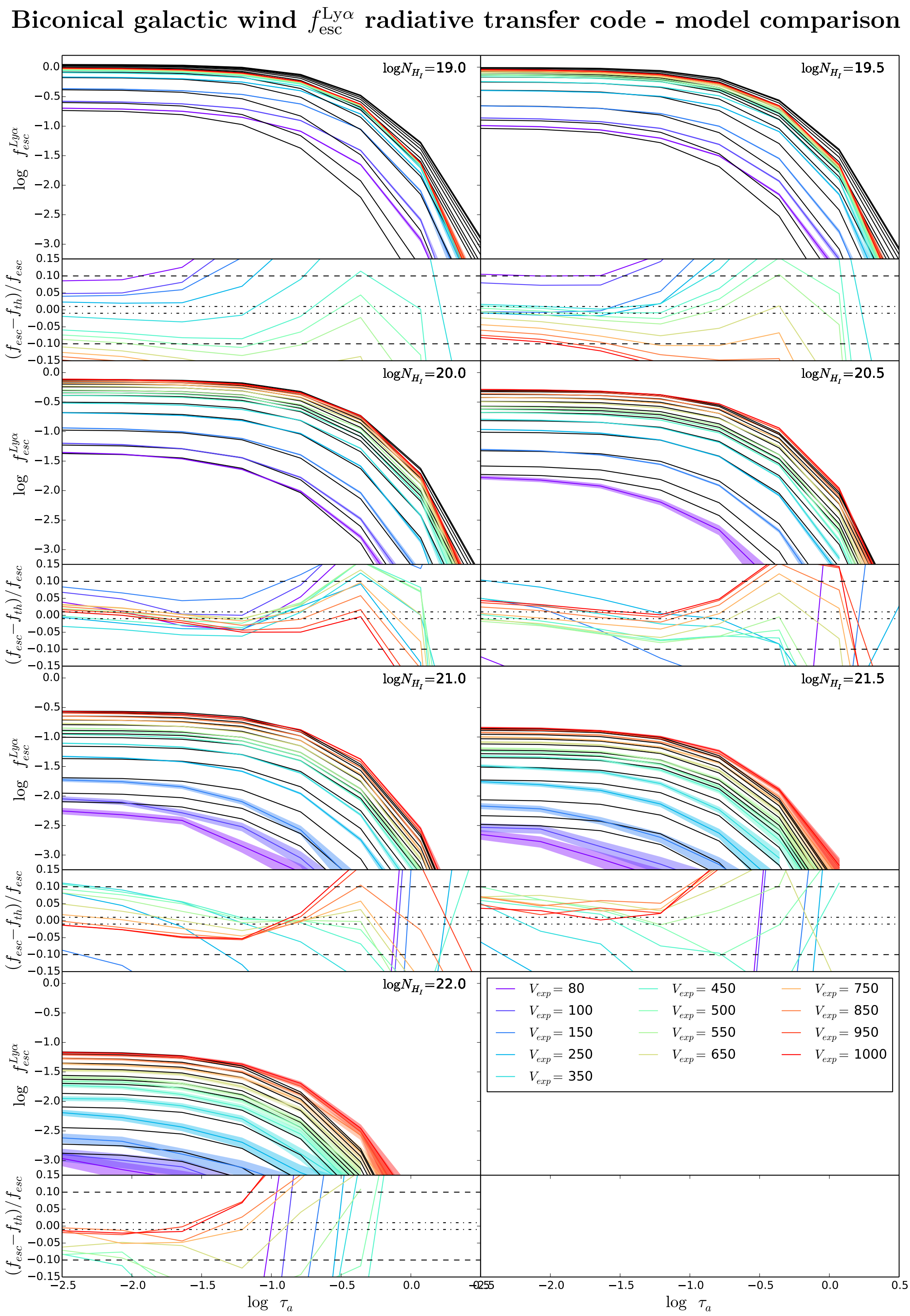


Fig. 2.14 Same as figure 2.14 but for the biconical galactic wind.

Table 2.5 Properties of the different mock catalogs and surveys.

Authors	z	$\mathrm{EW}_{0,\mathrm{cut}}[\text{Å}]$					$\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}\left[\mathrm{erg\ s}^{-1}\right]$				
		Survey	Thin Shell	Wind	Bicone	AM	Survey	Thin Shell	Wind	Bicone	AM
Kusakabe et al. (2018)	2.2	20.0	19.91	20.3	18.79	19.52	$1.62\ 10^{42}$	$1.54\ 10^{42}$	$1.92\ 10^{42}$	$1.38\ 10^{42}$	$1.6\ 10^{42}$
Bielby et al. (2016)	3.0	65.0	42.0	48.37	46.05	20.42	$1.62\ 10^{42}$	$1.33\ 10^{42}$	$1.57\ 10^{42}$	$1.47\ 10^{42}$	$1.48\ 10^{42}$
Ouchi et al. (2018)	5.7	20.0	20.06	20.06	18.77	21.45	$6.3\ 10^{42}$	$7.39\ 10^{42}$	$6.89\ 10^{42}$	$2.61\ 10^{42}$	$6.78\ 10^{42}$
Ouchi et al. (2018)	6.7	20.0	20.06	20.06	15.88	21.45	$7.9\ 10^{42}$	$7.7\ 10^{42}$	$6.11\ 10^{42}$	$1.86\ 10^{42}$	$8.37\ 10^{42}$

2.7.3 Choosing an EW and luminosity cut for the mock catalogues

In order to compare our clustering predictions with observations we construct mock catalogs that mimic the properties of several surveys at different redshifts. In general, there are several options for building mock catalogs to measure clustering.

The first one, for example, is to use the same selection criteria (flux depth, equivalent width cut, etc) than the observed samples. This first option is useful if all the properties used in the selection criteria are well reproduced by the models.

The LAE surveys studied in this work are limited by $\mathrm{L_{Ly\alpha}} > \mathrm{L_{Ly\alpha,cut}}$ and $\mathrm{EW_0} > \mathrm{EW_{0,cut}}$. In general $\mathrm{L_{Ly\alpha,cut}}$ and $\mathrm{EW_{0,cut}}$ are different for every survey. These values are listed in Table 2.5.

Our models are designed so they reproduce the abundance and luminosity distribution LAEs as we force them to fit, as good as possible, the observed LF at different redshifts. In detail, we combine different observations of the Lyα LF at the same redshift in order to calibrate our models. Because of this, the surveys that we use to study the clustering and calibrate our models, in general, use different selection criteria or the source sample is different. This could lead to discrepancies in the predicted number density of sources by our models imposing the clustering studies restrictions and the observed abundance of sources in these ones.

In particular, at $z = 2.2$ the survey constraining the clustering (Kusakabe et al., 2018) is, at least, partially included in one of the surveys used to calibrate the LF (Konno et al., 2016). Additionally, $\mathrm{EW_{0,cut}}$ is the same for all the surveys used to fit the LF (Cassata et al., 2011; Konno et al., 2016; Sobral et al., 2017a) and Kusakabe et al. (2018).

However, at $z = 3.0$ the selection criteria of the surveys used to fit the LF has different cuts in equivalent width (Cassata et al. (2011) $\mathrm{EW_{0,cut}} = 20$Å and Ouchi et al. (2008) $\mathrm{EW_{0,cut}} = 64$Å) while Bielby et al. (2016) (clustering measurements) has $\mathrm{EW_{0,cut}} = 65$Å.

The best scenario happens at redshifts 5.7 and 6.7, where the surveys used to calibrate our models (Ouchi et al., 2008; Konno et al., 2016) are practically the same in sky coverage and selection criteria than the ones used to constrain the clustering (Ouchi et al., 2010, 2018).

The second method to construct mock catalogs consists in matching the observed number density of sources. This can be achieve by relaxing the selection criteria. To minimize the possible secondary effects in the clustering due to changes in the selection criteria, we choose the combination that minimizes

$$\mathrm{Q} = (\log \mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{n}} - \log \mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{s}})^2 + (\log \mathrm{EW}_{0,\mathrm{n}} - \log \mathrm{EW}_{0,\mathrm{s}})^2, \tag{2.22}$$

where $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{s}}$ and $\mathrm{EW}_{0,\mathrm{s}}$ are the $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$ and $\mathrm{EW}_{0,\mathrm{cut}}$ imposed by each survey and $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{n}}$ and $\mathrm{EW}_{0,\mathrm{n}}$ define the iso-$\mathrm{n}_{\mathrm{LAE}}$ curve with the LAE observed abundance. In Table 2.5 we list $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{s}}$ and $\mathrm{EW}_{0,\mathrm{s}}$ for the different surveys and the used values of $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$ and $\mathrm{EW}_{0,\mathrm{cut}}$ to construct the mock catalogs.

In Figs. 2.15, 2.16, 2.17 and 2.18 we show the predicted $\mathrm{n}_{\mathrm{LAE}}$ by our different models for several $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$-$\mathrm{EW}_{0,\mathrm{cut}}$ combinations at $z = 2.2$, 3.0, 5.7 and 6.7 respectively. In these figures we also show $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$ and $\mathrm{EW}_{0,\mathrm{cut}}$ of each of the surveys used for clustering in black dashed lines. The intersection between these shows the location of the clustering surveys selection criteria. Additionally, it is shown the individual value of $\mathrm{n}_{\mathrm{LAE}}$ predicted by our models imposing the observational cuts (indicated with the white arrow). We also show the curve with constant $\mathrm{n}_{\mathrm{LAE}}$ matching the observed abundance (solid black line). Finally, the $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$-$\mathrm{EW}_{0,\mathrm{cut}}$ combination that minimize Eq. 2.22 is shown as a white dot.

At redshift 2.2 the predicted (using the survey selection criteria) and observed $\mathrm{n}_{\mathrm{LAE}}$ match quite well. Thus, $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$ and $\mathrm{EW}_{0,\mathrm{cut}}$ are very similar to $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{s}}$ and $\mathrm{EW}_{0,\mathrm{s}}$. However, the opposite case is found at $z = 3.0$, where predicted $\mathrm{n}_{\mathrm{LAE}}$ is heavily underestimated in comparison with observations. This is mainly due to the mismatch between the predicted EW_0 distribution and the observed one. This might be due to the difference in selection criteria used the authors of the works for constraining the LF and the work building the clustering sample. While $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$ is relatively similar to $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{s}}$, in order to recover the observed $\mathrm{n}_{\mathrm{LAE}}$, in all models, the value of $\mathrm{EW}_{0,\mathrm{cut}}$ is significantly lower than $\mathrm{EW}_{0,\mathrm{s}}$.

The scenarios at redshift 5.7 and 6.7 are quite similar. At both redshifts the predicted number density, using the survey selection criteria, and observed $\mathrm{n}_{\mathrm{LAE}}$ match quite well for the *Thin Shell*, *Wind* and AM-noRT samples. However, in the *Bicone* model $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$ and $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{s}}$ are very different. In particular, the *Bicone* model requires a low $L_{Ly\alpha,cut}$ in order to balance underestimation of abundance (see Fig. 2.3).

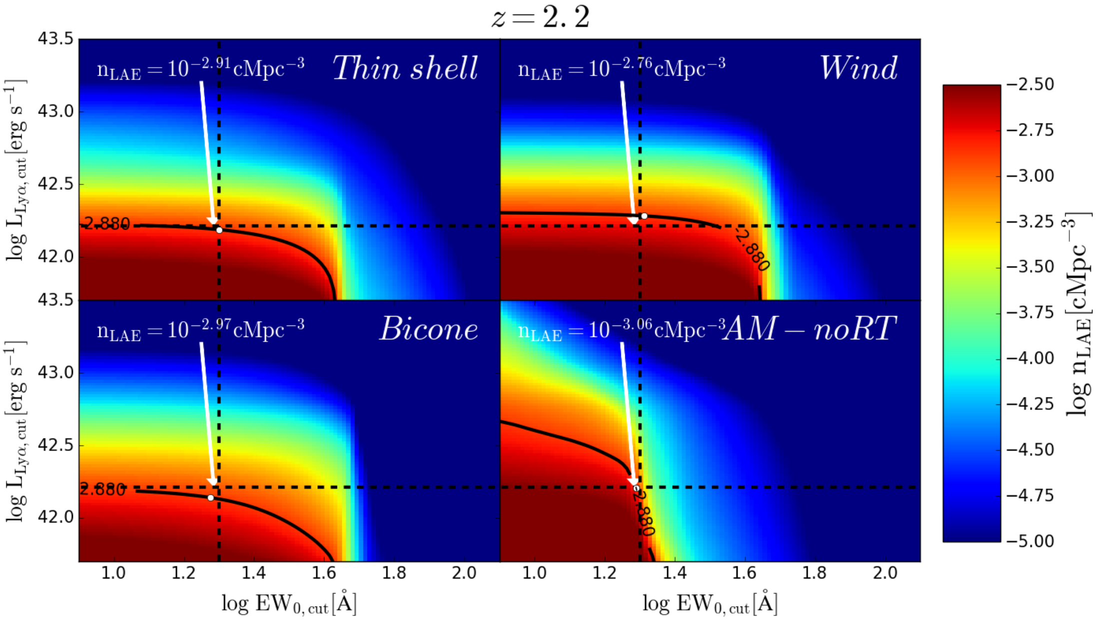


Fig. 2.15 Number density of LAEs n_{LAE} with Lyα luminosity $L_{Ly\alpha} > L_{Ly\alpha,cut}$ and Lyα rest frame equivalent width $EW_0 > EW_{0,cut}$ at redshift 2.2 for the *Thin Shell* (top left), *Wind* (top right), *Bicone* (bottom left) and AM-noRT (bottom right) model. In horizontal and vertical dashed black line we show the cut in $L_{Ly\alpha}$ and EW_0 respectively, in the survey at this redshift (Kusakabe et al., 2018). The place where these lines intersect sets the predicted n_{LAE} by our models which value is indicated in the same panel. The solid back line is the iso-number density curve of the observed n_{LAE}. The white dot indicates the position in the iso-number density curve that minimize the distance between our model prediction and the observed n_{LAE}.

.

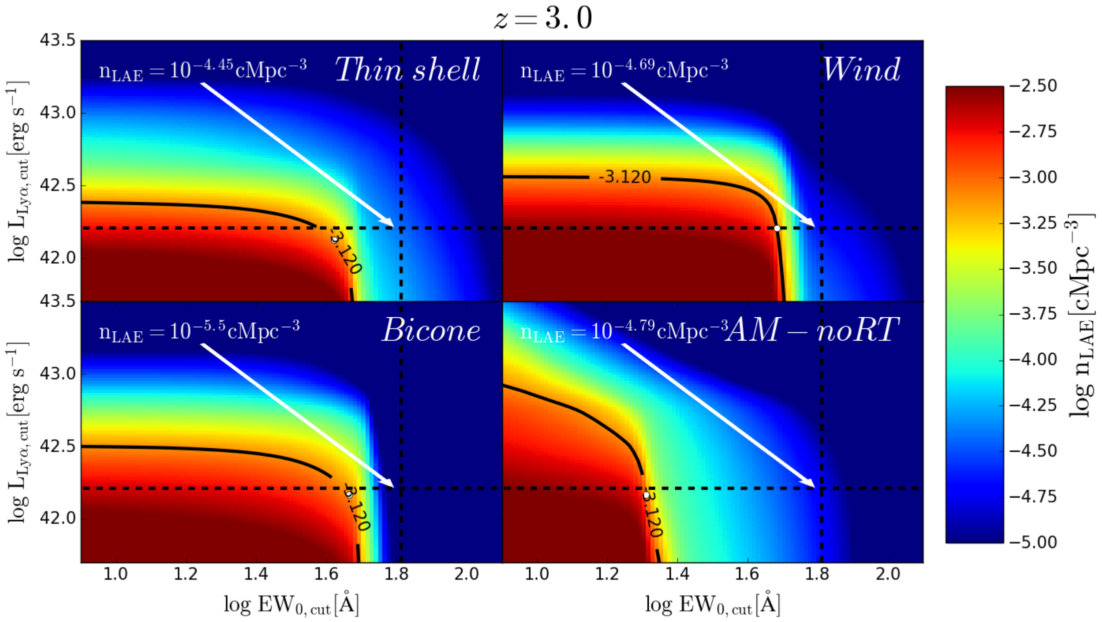


Fig. 2.16 Same as Fig. 2.15 but at redshift 3.0 (Bielby et al., 2016)

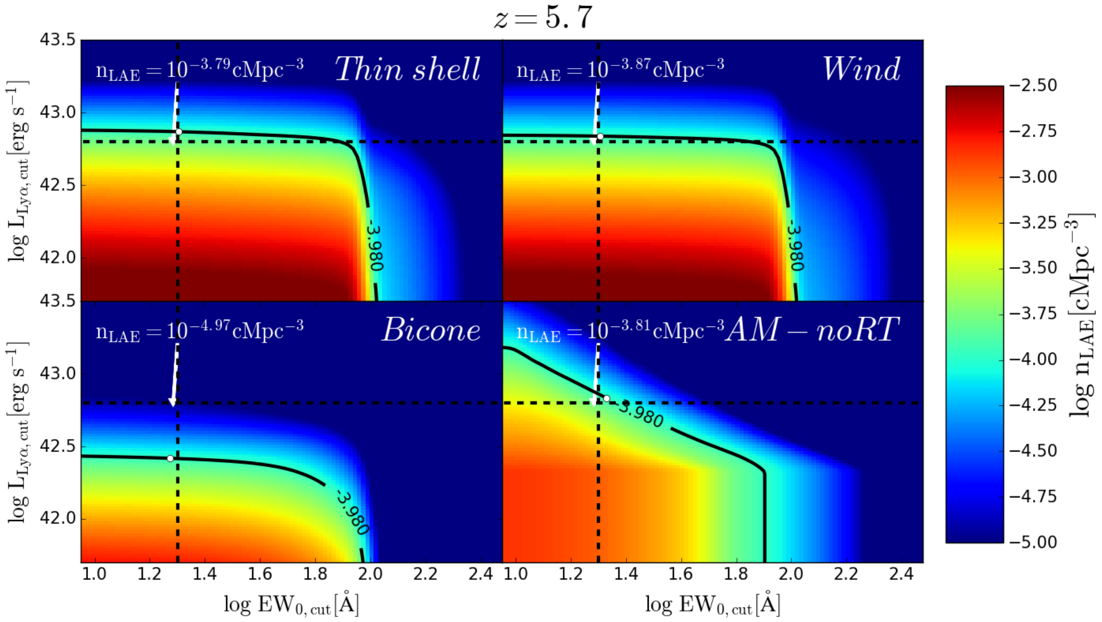


Fig. 2.17 Same as Fig. 2.15 but at redshift 5.7 (Ouchi et al., 2018)

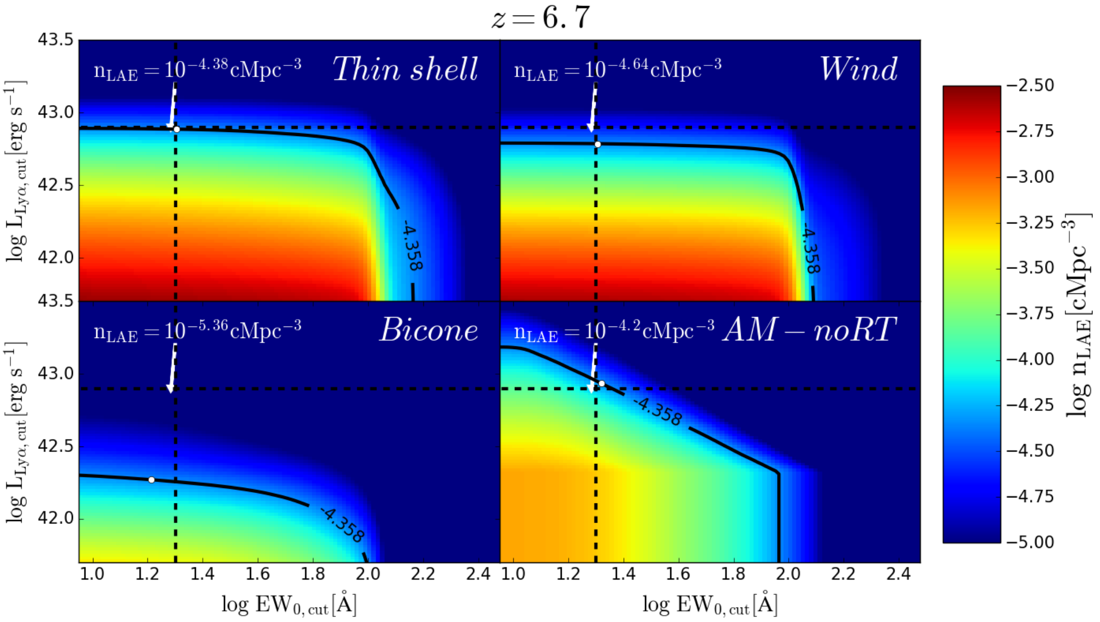


Fig. 2.18 Same as Fig. 2.15 but at redshift 6.7 (Ouchi et al., 2018)

Chapter 3

`FLaREON`: Fast Lyman-Alpha Radiative Escape from Outflowing Neutral gas

Abstract

We present `FLaREON` (Fast Lyman-Alpha Radiative Escape from Outflowing Neutral gas), a public `Python` package that delivers fast and accurate Lyα escape fractions and line profiles over a wide range of outflow geometries and properties. The code incorporates different algorithms, such as interpolation and machine learning to predict Lyα line properties from a pre-computed grid of outflow configurations based on the outputs of a Monte Carlo radiative transfer code. Here we describe the algorithm, discuss its performance and illustrate some of its many applications. Most notably, `FLaREON` can be used to infer the physical properties of the outflowing medium from an observed Lyα line profile, including the escape fraction, or it can be run over millions of objects in a galaxy formation model to simulate the escape of Lyα photons in a cosmological volume.

3.1 Introduction

Since the first evidence of star forming galaxies emitting Lyα photons (Steidel et al., 1996; Hu et al., 1998), more than two decades ago, observational campaigns targeting these sources have developed to become a standard technique to identify high redshift galaxies (e.g. Rhoads et al., 2000; Malhotra and Rhoads, 2002; Konno et al., 2016;

Sobral et al., 2017a; Ouchi et al., 2018). However, we are still far from a comprehensive understanding of this galaxy population due to the complex radiative transfer (RT) processes that Lyα photons experience (see Dijkstra, 2017, for a review).

The Lyα RT in astrophysical media can be addressed analytically for static, simplified geometries (e.g. Harrington, 1973; Neufeld, 1990). However, the limited validity of such approach encouraged the development of numerical Monte Carlo radiative transfer (MCRT) codes. In this approach, Lyα photons are tracked individually as they interact in arbitrarily complex 3D gas geometries. As a result, information about the fraction of photons that manage to escape, $f_{\rm esc}^{\rm Ly\alpha}$, and their resulting line profile is computed (Ahn et al., 2000; Zheng and Miralda-Escudé, 2002; Ahn, 2003; Verhamme et al., 2006; Gronke et al., 2016; Orsi et al., 2012).

One drawback of the MCRT technique is given by the time it takes to simulate an appropriate number of photons to obtain statistically significant results. The mean number of scattering events scales roughly proportionally with the Lyα rest-frame optical depth of the medium (Harrington, 1973). Hence, the computational time can vary by several orders of magnitude depending on the physical configuration probed. Typically, approximately $10^4 - 10^5$ photons are needed to retrieve the shape of the Lyα line profile with reasonable resolution. Such exercise becomes quickly prohibitively expensive when running MCRT codes for multiple configurations. One scenario where this requirement is needed is in Lyα line profile fitting (e.g. Mejias et al., in prep). Another example is to incorporate the Lyα properties of objects in a galaxy formation model run over a cosmological box (e.g. Orsi et al., 2012; Garel et al., 2012; Gurung López et al., 2018)

Here we address the problem described above by presenting `FLaREON`, a publicly available `python` package able to quickly predict multiple Lyman alpha line profiles and escape fractions with high accuracy. The basis of the results is a grid of configurations computed previously using the MCRT code `LyaRT` (Orsi et al., 2012).The outline of this work is as follows. In §2 we present `FLaREON`. In §3 we test its accuracy and in §4 we briefly explain how to exploit `FLaREON`. In §5 we illustrate some possible applications. Finally, we present our conclusions in §6.

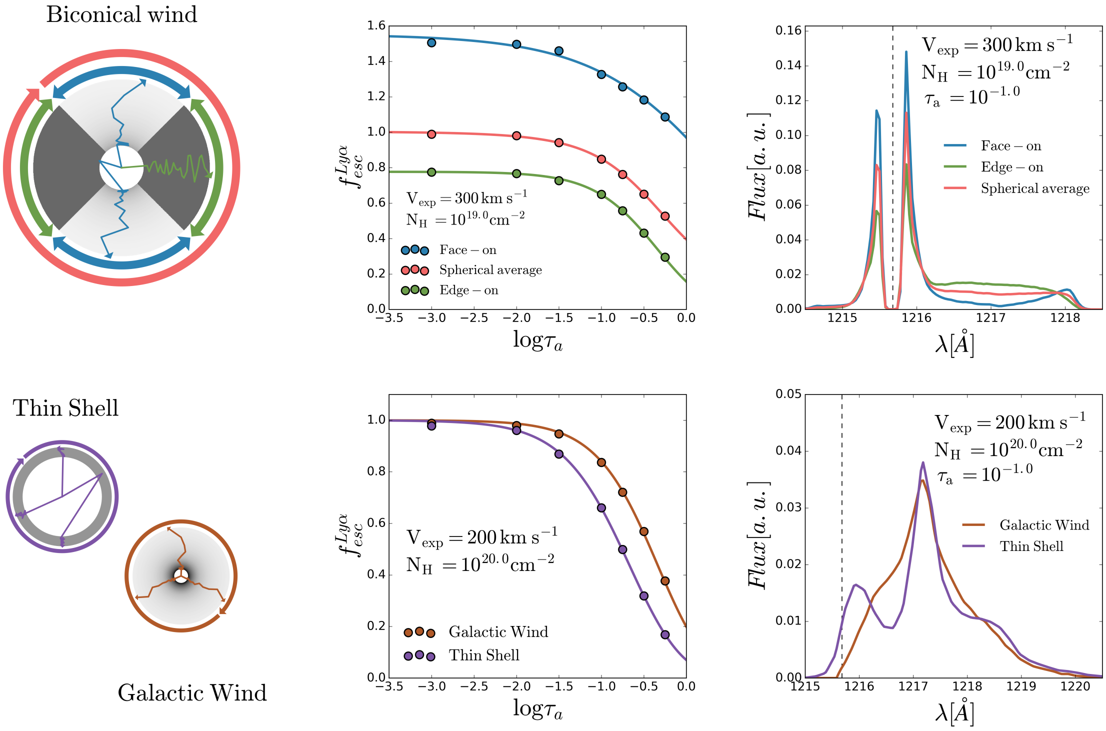


Fig. 3.1 The Lyα line properties in different geometries. The left panels show a cartoon of the Biconical Wind, Thin shell and Galactic Wind. The middle panels show the Lyα escape fraction as a function of dust optical depth of absorption τ_a for the three geometries and outflow configuration as shown in the legend. The right panels show examples of Lyα line profiles for the configurations described in the legend.

3.2 Code description

`FLaREON` [1] makes use of a grid of configurations run with the Monte Carlo radiative transfer code `LyaRT`[2] (Orsi et al., 2012). This grid covers three different outflow geometries and a wide range of gas properties. The goal of the code is to deliver a user-friendly public `python` package able to predict thousands of Lyman alpha line profiles and $f_{\rm esc}^{\rm Ly\alpha}$ in seconds with minimal user input.

3.2.1 Outflow Geometries

There are three different outflow geometries implemented in `FLaREON`: *Thin shell, Wind,* and *Biconical wind.* They all feature an empty inner cavity and an isotropic

---

[1] `https://github.com/sidgurun/FLaREON`

[2] `https://github.com/aaorsi/LyaRT`

monochromatic source of Lyα photons is placed in the centre. In all geometries, dust follows the gas density. The dust optical depth is defined as

$$\tau_a = (1 - A_{\mathrm{Ly}\alpha}) \frac{E_\odot}{Z_\odot} N_H Z, \tag{3.1}$$

where $E_\odot = 1.77 \times 10^{-21}\mathrm{cm}^{-2}$ is the ratio τ_a/N_H for solar metallicity, $A_{Ly\alpha} = 0.39$ is the albedo at the Lyα wavelength, $Z_\odot = 0.02$ (Granato et al., 2000), Z is the gas metallicity and $\mathrm{N_H}$ is the neutral hydrogen column density. We assume the temperature of the gas is constant at $T = 10000\ K$.

The Thin shell and Wind geometries are described in detail in Orsi et al. (2012) and Gurung López et al. (2018). In the following we briefly describe them. The Biconical wind described below is slightly different to that presented in Gurung López et al. (2018).

1. Thin shell: an isothermal uniform neutral hydrogen distributed in a thin layer with a radial expansion velocity $\mathrm{V_{exp}}$. This geometry has been widely used in the literature to study the escape of Lyα photons (e.g. Zheng and Miralda-Escudé, 2002; Ahn, 2004; Verhamme et al., 2006; Orsi et al., 2012).
2. Wind: a spherical isothermal distribution of neutral hydrogen with radial expansion velocity $\mathrm{V_{exp}}$ as implemented in Orsi et al. (2012). This geometry exhibits an empty spherical cavity with radius $\mathrm{R_{Wind}}$ (analog to $\mathrm{R_{inner}}$ in the thin shell) and a radially decreasing number density profile.
3. Biconical wind: this is a combination of an outflow with expanding wind geometry and a static isothermal uniform medium. The expanding wind outflow with $\mathrm{V_{exp,in}} > 0$ and $\mathrm{N_{H,in}}$ is confined in $\theta < \theta_{cone}$ and $\theta > \pi - \theta_{\mathrm{cone}}$, where θ_{cone} is measured from the polar axis. We define $\theta_{\mathrm{cone}} = \pi/4$. For $\theta_{\mathrm{cone}} < \theta < \pi - \theta_{\mathrm{cone}}$, the medium is static ($\mathrm{V_{exp,out}} = 0$) with column density $\mathrm{N_{H,out}} = f\mathrm{N_{H,in}}$. Here we arbitrary set $f = 10^3$. The column density of this geometry $\mathrm{N_H}$ is

   $$\mathrm{N_H} = (1 - \cos\theta_{\mathrm{cone}})\mathrm{N_{H,in}} + \cos\theta_{\mathrm{cone}}\mathrm{N_{H,out}} \tag{3.2}$$

   where $\mathrm{N_{H,in}}$ corresponds to the column density of the Wind geometry (see Gurung López et al., 2018).

Fig. 3.1 illustrates the main features of the three geometries described. For the Biconical wind, photons scattering through the dense and static torus are less likely to

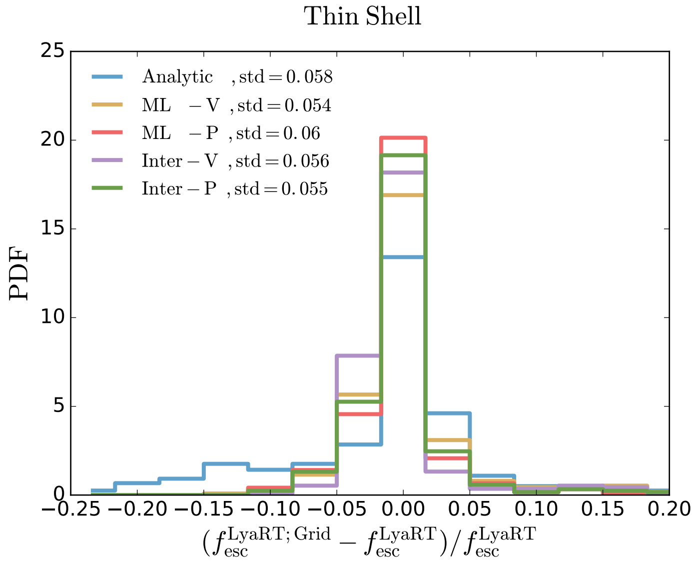


Fig. 3.2 Distribution of the relative difference between the output of `LyaRT` and the predictions of `FLaREON` for 300 random outflow configurations using the Thin Shell geometry. The result with the analytical expressions is shown in blue. The result with the machine learning algorithm using the outputs of `LyaRT` and using the fitting parameters (in equation 3.6) are plotted in yellow and red, respectively. Additionally, the algorithms using lineal interpolation and the output of `LyaRT` and using the fitting parameters are shown in purple and green respectively. The legend displays the standard deviation of each `FLaREON` algorithm.

escape compared to those traveling through the thin outflow. The dipolar nature of the scattering of Lyα photons through a neutral hydrogen medium makes them likely to back-scatter within the inner cavity until they escape through the thin medium. For $\theta_{\rm cone} = \pi/4$, $\sim$ 70% of the photons are emitted towards the thick torus and only a few ($\sim$ 30%) towards the bicone.

The inclination of the Biconical wind with respect to the observer leads to different escape fractions. This is shown in the middle panel of Fig. 3.1 by computing the escape fraction $f_{\rm esc}^{\rm Ly\alpha}$ of Lyα photons though the bicone (face-on) or through the torus (edge-on). The escape fraction is defined as the ratio between the number of photons that escape through a given direction over the number of photons emitted towards that direction. The three cases shown in Fig. 3.1 (edge-on, face-on and the spherical average) show significant differences. If the geometry is observed face-on, $f_{\rm esc}^{\rm Ly\alpha}$ reaches values greater than 1, while if it is edge-on $f_{\rm esc}^{\rm Ly\alpha}< 1$ even if there is no dust. This is caused by the large optical depth of the torus that beams the Lyα photons towards the bi-

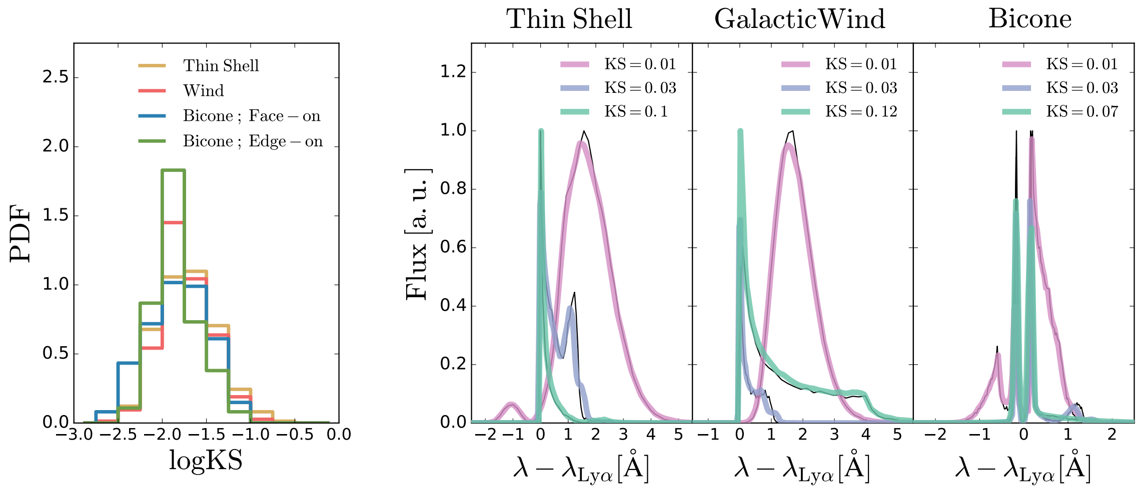


Fig. 3.3 **Left** : Normalized distribution of the KS statistic of the Lyα line profile computed by LyaRT and predicted by `FLaREON` in the same 300 random $\{\mathrm{V_{exp}},\, \mathrm{N_H},\, \tau_a\}$ combinations for the Thin Shell (yellow), the Wind (red), and the bicone face-on (blue) and edge-on (green). **Right** : comparison between the `LyaRT` output (solid black lines) and `FLaREON` predictions (colored solid lines) in some of the 300 random configurations. We show Thin Shell, Wind and Bicone configurations from left to right. Additionally, in the legend indicates the KS between the `LyaRT` and `FLaREON` line profiles.

cone. In the spherical average $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ can reach a value up to 1 for dust-free configurations.

The right panel of Fig. 3.1 shows the resulting line profiles for the bicone at the three different orientations. The differences in the line profile between the three cases reflect the different scattering histories of photons escaping through the outflow or the thick torus.

Fig. 3.1 also shows the corresponding $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ and line profiles for the Thin shell and Wind, respectively. Here, there is no difference in the escape of Lyα photons due to the orientation of the outflows.

3.2.2 Monte Carlo configuration grids

`FLaREON` is based in the outputs of `LyaRT` in a grid of configurations spanning a wide range of phyisical properties. We build two grids for each outflow geometry; one to infer $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ and another to predict Lyα line profiles. For the $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$, the grids are constructed using a number of photons $N_p = 10^4$. Hence, the lowest value of $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ computed is 10^{-4}.

For the line profile grids, we use instead $N_p = 10^5$, since we need more photons to fully recover the shape of the resulting line profiles. In order to speed-up the computational time in the latter case, we implement an acceleration procedure to dismiss scattering events that result in frequency changes below a critical value of $x_{\rm crit} = 3$, where x is the frequency of photons in Doppler units (see, e.g. Dijkstra et al., 2006; Laursen and Sommer-Larsen, 2007; Orsi et al., 2012). Those scattering events have no significant impact on the resulting line profile, but skipping them can improve the performance of the Monte Carlo calculation by orders of magnitude. Since the bicone geometry is split into two lines of sight we increased the number of photons to 10^5 and 10^6 for the $f_{\rm esc}^{{\rm Ly}\alpha}$ and line profile grids, respectively.

The $\rm V_{exp}$-$\log \rm N_H$ parameter space covered in the grids is defined as follows:

$$\rm V_{exp}[km\ s^{-1}] = \begin{cases} [10\ ,100]\ , & \Delta V_{exp} = 10\ km\ s^{-1} \\ [100\ ,1000]\ , & \Delta V_{exp} = 50\ km\ s^{-1} \end{cases} \tag{3.3}$$

where $\Delta \rm V_{exp}$ is the step between evaluations. Additionally, the neutral hydrogen column density N_H is mapped in bins of $\Delta \log \rm N_H = 0.25$ and spans the following range:

$$\log \rm N_H[cm^{-2}] = [17\ ,22]\quad ,\quad \Delta\log N_H[cm^{-2}] = 0.25 \tag{3.4}$$

The values of dust optical depth τ_a where the grid is sampled are different for the $f_{\rm esc}^{{\rm Ly}\alpha}$ and line profile grids. We have checked that a sparse sampling of $f_{\rm esc}^{{\rm Ly}\alpha}$ is sufficient to deliver good results. These values are $\log \tau_a = [-3.0\ , -2.0\ , -1.5\ , -1.0\ , -0.75\ , -0.5\ , -0.25\ ,\ 0.0]$. The line profile grids, on the other hand, spans a wider range and more frequently in τ_a to track properly the evolution with dust optical depth of the Lyα line profile. We cover the low dust range with higher density as we found that the evolution is stronger in this range. Finally, the dust optical depth values are

$$\log \tau_a = \begin{cases} [-3.75\ , -1.500]\ , & \Delta \log \tau_a = 0.25 \\ [-1.50\ , -0.125]\ , & \Delta \log \tau_a = 0.125 \end{cases} \tag{3.5}$$

where $\Delta \log \tau_a$ is the step between evaluations.

All together, the total number of configurations sampled for each geometry is 4704 and 12348 for the $f_{\rm esc}^{\rm Ly\alpha}$ and line profile grids, respectively.

3.2.3 Predicting the Lyman α properties.

`FLaREON` allows the user to choose among three different methods to compute the $f_{\rm esc}^{\rm Ly\alpha}$from the outflow properties:

1. **The first method** uses directly the values of $f_{\rm esc}^{\rm Ly\alpha}$ given by `LyaRT` to build a multi-dimensional linear interpolation grid or to train a machine learning algorithm. For the latter, `FLaREON` incorporates 'extra trees', 'random forest' and 'k-nearest neighbors', using the `python` module `scikit-learn` (Pedregosa et al., 2011).

2. **The second method** consists in using a parametric equation that links $f_{\rm esc}^{\rm Ly\alpha}$ and one or several gas properties. In particular, we fit the $f_{\rm esc}^{\rm Ly\alpha}$computed by `LyaRT` as a function of the dust optical depth in each node of the $\rm V_{exp}$-$\log \rm N_H$ space to the function

$$f_{\rm esc}^{\rm Ly\alpha} = k_3 \left[\cosh \sqrt{k_1 \tau_a^{k_2}} \right]^{-1} , \tag{3.6}$$

   where k_1, k_2 and k_3 are free parameters. Note that $k_3 = 1$ in the Thin shell and Wind geometries since if there is no dust $f_{\rm esc}^{\rm Ly\alpha} = 1$. However, in the Bicone geometry, as we divide in edge-on and face-on, in general, $k_3 \neq 1$ as discussed in §3.2.1. The best fitting parameters are then used to build the lineal interpolated grid or to train the machine learning algorithms in this mode.

3. **The third method** consists on an updated version of the analytic $f_{\rm esc}^{\rm Ly\alpha}$ expressions described in Gurung López et al. (2018). Those $f_{\rm esc}^{\rm Ly\alpha}$ parametric equations have been recalibrated for the Thin Shell and Wind in our grid, which is denser and wider in $\rm V_{exp}$ and $\rm N_H$. Due to the complexity of $f_{\rm esc}^{\rm Ly\alpha}$ in the face-on and edge-on configurations, `FLaREON` does not include any analytic expression for the biconical outflow.

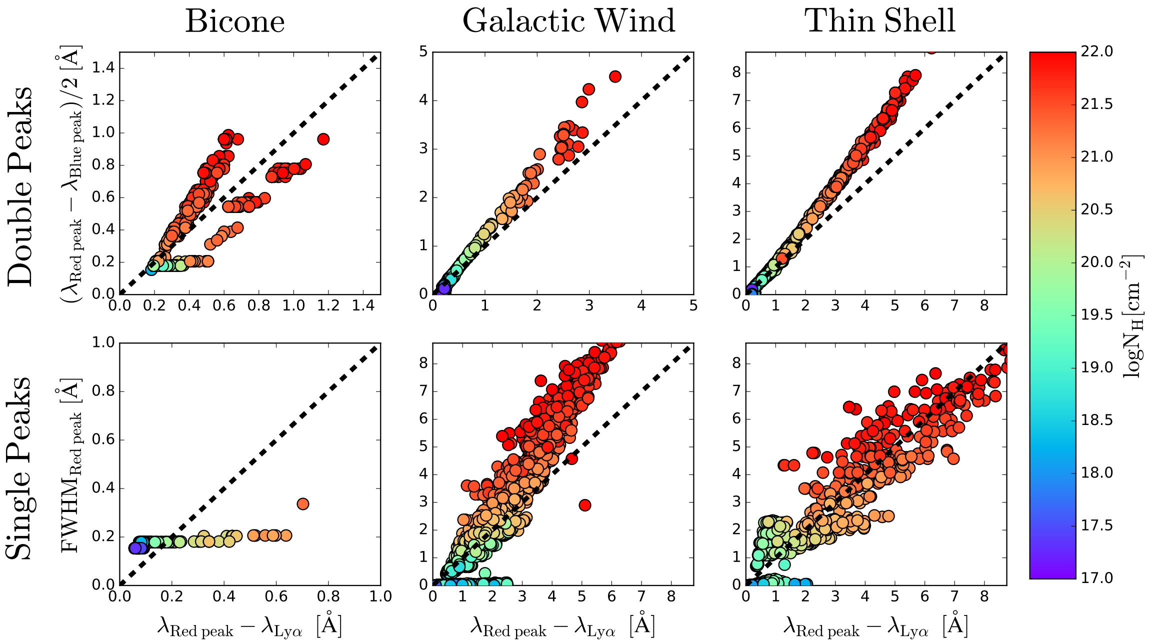


Fig. 3.4 Correlation between some line profile properties produced by the Bicone, Galactic Wind and Thin Shell geometries from left to right. **Top** : shift of the red peak from Lyα wavelength as a function of half of the separations between the red and the blue peaks in the line profile. Only configurations producing double peak profiles were taken into account for this panel. **Bottom** : shift of the red peak from Lyα wavelength as a function of the full width half maximum of the red peak. Only configurations producing line profiles without a blue peak were taken into account for this panel. In every panel, the $\mathrm{N_H}$ is color coded. The one-to-one relation is plotted in black dashed lines.

To address the problem of predicting Lyα line profiles, analytic expressions and machine learning algorithms perform poorly compared to a multi-dimensional linear interpolation of the `LyaRT` outputs. Hence, unlike the case of the $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$, the Ly$\alpha$ line profiles are computed using only a multi-linear interpolation of the grids.

3.3 Validation of the code

Here we compare the performance of different methods to obtain $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$. We compute 300 random values for $\mathrm{V_{exp}}$, $\log \mathrm{N_H}$ and $\log \tau_a$ using a Latin hypercube algorithm to populate randomly and homogeneously the three dimensional space within the range covered by the grids. Fig. 3.2 shows the relative difference between the `FLaREON` predicted $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ and the `LyaRT` output in those 300 random configurations using the

Thin Shell geometry. We find a remarkably good match between `LyaRT` and `FLaREON`. The analytic functional form has worse performance (although $\sim 70\%$ is above 90% accuracy) since these were optimized to a smaller $V_{\rm exp}$-$N_{\rm H}$ region. Additionally, $\sim 80\%$ of the configurations using directly $f_{\rm esc}^{\rm Ly\alpha}$ from `LyaRT` to train or to interpolate have an accuracy better than 90%. The method that gives the best results is the parametric interpolation, as $\sim 95\%$ and $\sim 50\%$ of the configurations have relative differences below 0.1 and 0.01. Other outflow geometries perform similarly.

To quantify the performance of `FLaREON` predicting Lyα line profiles we perform a Kolmogorov–Smirnov (KS) test over the 300 random outflow configurations for each geometry. Fig.3.3 shows the resulting KS distributions of the 300 random configurations for all the geometries. We find a very good agreement between the Lyα line profiles computed by `LyaRT` and those predicted by `FLaREON`. For all geometries the KS distribution peaks around $\mathrm{KS} = 10^{-2}$, implying that the typical maximum difference of the cumulative line profiles is $\approx 1\%$ of the flux. Additionally, about 90% of the random samples exhibit $\mathrm{KS} < 0.05$ for all geometries.

The right panels in Fig. 3.3 show some example geometries where the outputs of Monte Carlo `LyaRT` and `FLaREON` are compared. Overall, the differences between the Monte Carlo code and `FLaREON` are negligible.

3.4 Hands on FLaREON

In this section we ilustrate how to execute `FLaREON`. After the installation, running `FLaREON` should only take a few `python` command lines.

A simple script to compute the Lyα escape fraction and line profiles for some given Thin Shell configurations is given below:

```
import FLaREON as Lya, numpy as np, pylab as plt
# 1) We define the configuration parameters.
# 1.1) Expansion velocity in km/s :
V_exp = [ 50 , 100 , 200 , 300 ]
# 1.2) Logarithm of column densities in cm**-2 :
log_NH= [ 18 , 19 , 20 , 21 ]
```

```
# 1.3) Dust optical depths :
tau_a  = [ 0.5 , 0.1 , 0.05 , 0.01 ]
# 1.4) Select a geometry :
Geometry = 'Thin_Shell'
# 2) Compute the escape fractions.
f_esc = Lya.RT_f_esc( Geometry, V_exp ,
                      log_NH  , tau_a )
# 3) Compute the line profiles.
# 3.1) Define the wavelength range in meters.
wavelength = np.linspace( 1213e-10, 1224e-10, 1000 )
# 3.2) Execute FLaREON
lines = Lya.RT_Line_Profile( Geometry, wavelength,
                             V_exp   , log_NH    ,
                             tau_a               )
# 4) Display the line profiles.
for line in lines :
    plt.plot( wavelength , line )
plt.show()
```

Other examples can be found in the GitHub repository, including the coupling of `FLaREON` with other popular public codes such as `emcee`[3] (Foreman-Mackey et al., 2013a) to perform Lyα line fitting.

3.5 Some applications

In this section we present a small glimpse of the potential scientific applications of `FLaREON`.

3.5.1 Line profile properties

Recently, Verhamme et al. (2018) characterized the Lyα line profiles to infer their displacement redwards of the line centre, allowing to infer the systemic redshift of a source from the Lyα line only. This is done by measuring the difference in wavelength

---

[3] http://dfm.io/emcee/current/

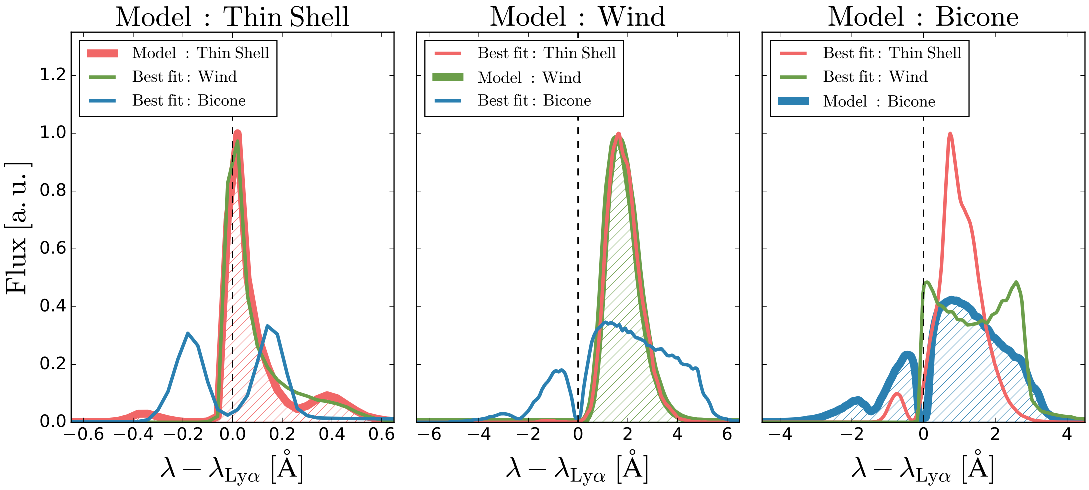


Fig. 3.5 Comparison of the randomly selected model's Lyα line profiles and the best fitting outflow configurations using different gas geometries. The Thin Shell, Galactic Wind and Bicone are shown in red, green and blue respectively. Additionally line profile used as model is highlighted by thicker lines. The model outflow geometry changes from Thin Shell, to Galactic Wind to Bicone from left to right.

between different peaks, if there is more than one, or by measuring the FWHM of a single peak.

Fig. 3.4 shows the relation between different line profile properties for 1000 different configurations spanning the full $V_{\rm exp} - N_H - \tau_a$ space using the three geometries. In the top panels we show the relation between the red peak shift and half of the distance between the red and blue peak in only configurations exhibiting both peaks. In general, `FLaREON` predicts a tight correlation between these properties and the neutral hydrogen column density. We find that `FLaREON` reproduce the correlation found by Verhamme et al. (2018) for shift of the read peak smaller than 2Å where trend leaves the one-to-one relation and the slope increases.

In the bottom panels of Fig.3.4 we show the relation of the shift and the FWHM of the red peak in configurations where only the red peak was found. Overall, `FLaREON` also predicts a correlation between these properties and the column density, although with a greater scatter. Our results with `FLaREON` are consistent with the findings of Verhamme et al. (2018) for the Thin shell and Galactic Wind geometry.

Table 3.1 Description of randomly selected configurations (*models*) and their best fitting configuration in other outflow geometries.

Configurations	V_{exp} [km/s]	log N_{H} [cm^{-2}]	log τ_{a}	log KS	f_{esc}
Model : Thin Shell	55.0	17.62	-0.23	-	0.43
Best fit : Wind	62.4	17.14	-0.25	-1.5	0.58
Best fit : Bicone	890.5	17.11	-0.13	-0.5	0.46
Model : Wind	123.0	20.25	-1.29	-	0.88
Best fit : Thin Shell	93.2	20.5	-0.29	-2.2	0.07
Best fit : Bicone	616.3	22.0	-1.38	-0.6	0.8
Model : Bicone	375.0	21.42	-2.37	-	0.97
Best fit : Thin Shell	93.0	19.64	-2.75	-0.8	0.99
Best fit : Wind	348.2	19.5	-1.75	-0.7	0.98

3.5.2 Extract outflow information from Lyα line profiles.

One of the most attractive application of `FLaREON` consists in inferring outflow properties from measured Lyα line profiles. Usually, in this kind of analysis only one outflow geometry is implemented (e.g. Orlitová et al., 2018a; Gronke, 2017). However, since `FLaREON` includes several gas configurations the analysis can be extended to different outflow geometries. In this section we give a glimpse of the advantages of using several geometries and we will exploit further this idea in an upcoming work (Mejias et. al. , in prep).

To study how the inferred outflow properties depend on the gas geometry we start by generating a Lyα line profile with a given gas geometry and random outflow parameters. We refer to this as *model*. Then, we perform an MCMC analysis combining `FLaREON` and `emcee` (Foreman-Mackey et al., 2013a) with the other gas geometries to fit the line profile of the *model*.

We performed this test for three *models*, each with a different gas geometry. For the biconical geometry we use the spherically averaged configuration, but we check that the results resemble that obtained with the edge-on and face-on configurations. These few examples illustrate the degeneracy between gas geometry and outflow parameters.

In Fig. 3.5 we show a comparison between the *model*'s Lyα line profile and the best fits of the other outflow geometries, whereas in Table 3.1 we summarize the results of the MCMC. On one hand, the morphology of the Lyα line profiles produced by the Bicone are very different ($KS > 0.1$) from the line profiles generated by the Wind and Thin Shell. Hence, there is not confusion between these gas geometries. On the other hand, the Thin Shell and Wind line profiles resemble and fit each other Lyα line profile with very good agreement ($KS < 0.1$). However, since the gas morphology impacts the resulting line profile, the inferred properties from the fit, generally, do not match the *model*'s characteristic.

Additionally, the $f_{\rm esc}^{\rm Ly\alpha}$ computed from the inferred outflow properties differs from the $f_{\rm esc}^{\rm Ly\alpha}$ of the *model*. This increases the difficulty of calculating the intrinsic Lyα flux emitted before the radiative transfer processes by using only Lyα emission.

3.6 Conclusions

In this work we have introduced `FLaREON`, a user-friendly public `python` code based on the radiative transfer Monte Carlo code `LyaRT` (Orsi et al., 2012). This code is able to predict Lyα line profiles and Lyα escape fractions for different outflow geometries in a wide range of outflow properties without the need to run a Monte Carlo code.

`FLaREON` includes three different outflow geometries, an expanding Thin shell (e.g. Verhamme et al., 2006; Orsi et al., 2012; Gurung López et al., 2018), galactic wind (Orsi et al., 2012; Gurung López et al., 2018), and a biconical outflow surrounded by a very thick static torus.

In order to predict the Lyα line profile and $f_{\rm esc}^{\rm Ly\alpha}$ `FLaREON` interpolates or makes use of machine learning algorithms over previously-computed grids of configurations. This grids are very dense and cover a wide range of $\rm V_{exp}$, $\rm N_H$ and τ_a. These are composed by 4704 and 12348 different outflow configurations to predict the $f_{\rm esc}^{\rm Ly\alpha}$ and line profile, respectively.

We have analyzed the performance of `FLaREON` of the different geometries and predicting method implemented. `FLaREON` is able to predict thousands of $f_{\rm esc}^{\rm Ly\alpha}$and line profiles with remarkably high accuracy. The error in $f_{\rm esc}^{\rm Ly\alpha}$ is typically below 5%. Additionally, the Kolmogorov-Smirnov test delivers values about 0.01 between the fully computed `LyaRT` line profiles and `FLaREON` predictions.

In future works we plan to exploit the `FLaREON` capabilities to predict thousands of line profiles and escape fractions to extract outflow physical information such as $\rm V_{exp}$ or $\rm N_H$ from observed spectra and to populate large cosmological volumes with Lyα emitters.

Chapter 4

The impact of the intergalactic medium in Lyα selected samples

Abstract

In the near future galaxy surveys will target Lyman alpha emitting galaxies (LAEs) to unveil the nature of the dark energy. During the last years it has been suggested that the observability of LAEs is coupled to the large scale properties of the intergalactic medium. Such coupling could introduce distortions in the observed clustering of LAEs, adding a new potential difficulty to the interpretation of upcoming surveys. We present a model of LAEs that incorporates Lyα radiative transfer processes in the interstellar and intergalactic medium. The model is implemented on the `GALFORM` semi-analytic model of galaxy of formation and evolution. We find that the radiative transfer inside galaxies produces selection effects over galaxy properties. In particular, observed LAEs tend to have low metallicities and intermediate star formation rates. At low redshift we find no evidence of correlation between the spatial distribution of LAEs and the intergalactic medium properties. However, at high redshift the LAEs are linked to the line of sight velocity and density gradient of the intergalactic medium. The strength of the coupling depends on the outflow properties of the galaxies and redshift. This effect modifies the clustering of LAEs on large scales, adding non linear features. In particular, our model predicts modifications in the shape and position of the baryon acoustic oscillation peak. This work highlights the importance of including radiative transfer physics in the cosmological analysis of LAEs.

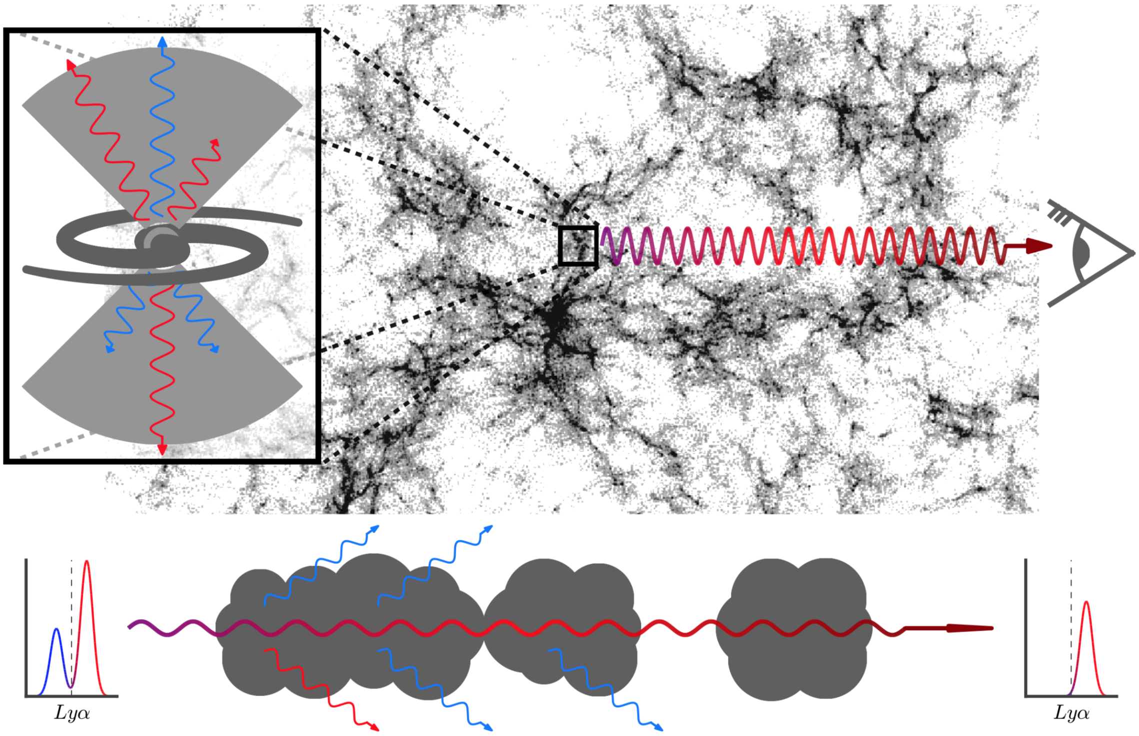


Fig. 4.1 Illustration of the journey of Lyα photons since they are emitted until they reach the observer. Since Lyα photons are generated in star forming regions they go through RT processes until escaping through galaxy outflows. Then they enter into the IGM, where photons interacting with H_I are scattered out of the line of sight and never reach the observer.

4.1 Introduction

Galaxies exhibiting strong Lyman-α emission, the so-called LAEs, are one of the most important tracers of the high redshift Universe in modern astrophysics. After their first detection about 20 years ago (e.g. Steidel et al., 1996; Hu et al., 1998; Rhoads et al., 2000; Malhotra and Rhoads, 2002) studies have focused on finding LAEs at low (Orlitová et al., 2018b; Henry et al., 2018) and high redshift (Ouchi et al., 2008; Oyarzún et al., 2017; Matthee et al., 2017; Caruana et al., 2018) including at the epoch of reionization (Sobral et al., 2015; Ouchi et al., 2018; Shibuya et al., 2018).

In the coming years, cosmological surveys such as HETDEX (Hill et al., 2008b) and J-PAS (Benitez et al., 2014b) aim to unveil the mystery of the dark energy. These surveys will scan the sky chasing LAEs to trace the underlying dark matter density fluctuations and make clustering measurements. Additionally, these surveys will contribute notably to the knowledge of galaxy formation and evolution. Thus, it is becoming timely to understand properly the selection function of LAEs and how it might affect the apparent spatial distribution and galaxy properties.

The large cross-section of neutral hydrogen atoms around Lyα wavelengths makes these photons suffer multiple scattering events that modify their frequency and direction. These frequency changes produce the characteristic Lyα line profiles widely studied (Verhamme et al., 2006; Gronke et al., 2016). The typical distance covered by Lyα photons inside atomic hydrogen is increased drastically due to multiple scattering events, making these photons very sensitive to dust absorption.

In order to understand the complex escape of Lyα photons consider the illustration shown in Fig.4.1. The journey of Lyα photons begins when a star formation episode takes place inside a galaxy. During these events, hot massive stars (mostly *O-type* and *B-type*) emit high energy photons capable of ionizing neutral hydrogen. A fraction of these photons dissociates the $\mathrm{H_I}$ in the surroundings of the star forming region. Another part is absorbed by dust grains, while the rest escapes from the galaxy and ionizes $\mathrm{H_I}$ in the intergalactic medium.

Then, the free electrons within the ionized region surrounding the young stars recombine with $\mathrm{H_{II}}$ ions in excited energy levels, causing a cascade until electrons reach the ground energy level. Lyα photons are emitted when an electron decays from the first excited level to the fundamental level, an even that occurs with probability $\sim 2/3$ (Spitzer, 1978)) per ionizing photon.

Lyα photons have to get through the intricate interstellar medium (ISM) before escaping the galaxy (Fig. 4.1 top left panel)(Neufeld, 1991). The ISM morphology includes dusty gas rich regions such as bars (Spinoso et al., 2017), arms (Kormendy, 2013), $\mathrm{H_{II}}$ bubbles, outflows (Cazzoli et al., 2016) and other structures that complicate the radiative transfer of Lyα photons. The resonant scattering inside the ISM enhances dust absorption and causes that only a fraction of the emitted Lyα photons manages to escape from the galaxy. Additionally, in this process the Lyα line profile is modified due to consecutive $\mathrm{H_I}$ interactions (Harrington, 1973). The final Lyα escape fraction and line profiles depend strongly on the ISM topology and kinematics (Verhamme et al., 2006; Gurung-Lopez et al., 2018).

After emerging from galaxies (Fig.4.1 zoom panel), Lyα photons enter into the IGM and interact with the H_I atoms within, producing further Lyα scattering events (Fig.4.1 main panel). While inside galaxies the flux reduction is due to dust absorption, in the IGM the Lyα photons are scattered out of the line of sight, as illustrated in the bottom panel of Fig.4.1. The Hubble flow redshifts the emitted photons, causing further IGM absorptions at wavelengths bluewards of Lyα while redder photons travel freely.

The RT processes, occurring in both the ISM and IGM, complicate the selection function of galaxy surveys using Lyα line detection as a tracer. Galaxies selected with this technique tend to have low metallicity and high specific star formation rates (Sobral et al., 2018a). Observational evidence also suggests that this galaxy population lies in low density environments (Shimakawa et al., 2017). Further observations are needed to cast light on this matter.

The RT in the ISM has been modeled and explored using Monte Carlo Radiative Transfer codes (Zheng and Miralda-Escudé, 2002; Ahn, 2003; Verhamme et al., 2006; Orsi et al., 2012; Gronke et al., 2016; Gurung-Lopez et al., 2018). These tools generate photons in H_I structures and track the subsequent interactions, changes in direction, frequency and possible absorptions. Monte Carlo Radiative Transfer codes have been implemented in cosmological simulations in the ΛCDM scenario to understand the effect of the ISM on the selection function of LAEs (Orsi et al., 2012; Garel et al., 2012; Gurung López et al., 2018).

The radiative transfer inside the IGM has been implemented in several ways. For example, Dijkstra et al. (2007) made use of analytic expressions to determine the velocity field, density field and ionization state of the IGM around galaxies as a function of some galaxy properties like the host halo mass or the circular velocity. Meanwhile, Laursen et al. (2011a) studied the IGM transmission around the Lyα wavelength in a hydrodynamic simulation tracking Lyα rays along different lines of sight. A different approach was taken by Zheng et al. (2010), where they run a Monte Carlo radiative transfer code (similar to the algorithms used to model the ISM) to study the observability of LAEs as a function of the IGM large scale properties. In particular, Zheng et al. (2010) found that the resonant nature of Lyα might introduce new clustering features in galaxy samples selected by Lyα detection. However, using a higher resolution simulation, Behrens et al. (2017) claimed to find only a marginal coupling between LAEs and the IGM.

If the IGM impacts significantly on the observed spatial distribution of LAEs, this could introduce dramatic biases into the cosmological interpretation of clustering data from surveys like HETDEX (Hill et al., 2008a), and future space mission concepts, such as the ATLAS Probe (Wang et al., 2018b) and the Cosmic Dawn Intensity Mapper (Cooray et al., 2016). Hence, understanding the role of the IGM in shaping the observed properties of high redshift LAEs is of crucial importance.

This is the second of a series of papers that tackle the selection effects on LAE caused by the RT of Lyα resonant scattering nature. In our first paper (Gurung López et al., 2018) we focused in the RT inside the ISM and how it determines the properties of galaxies observed as LAEs. Here, we expand our model and include the RT in the nearby IGM. We study the coupling between Lyα observability and different IGM large scale properties and how this modifies the clustering of LAEs.

This work is structured as follows : In §4.2 we present our model and its calibration (§4.3). Then, in §4.4 we briefly study the selection function of galactic properties while in §4.5 we focus on the selection effects of the IGM and its impact on LAE clustering (§4.6). Finally, we compare our work with the literature (§4.7) and present our conclusions (§4.8).

4.2 A comprehensive model for LAEs.

The model presented here is a follow up of the work presented in Gurung López et al. (2018). We combine a wide range of physical scenarios in order to assemble a realistic LAE model. Our model is built upon four main pillars:

1. The `P-Millennium` N-Body simulation (Baugh et al., 2019), a stare-of-the-art dark matter N-body simulation with box size $(542.16\mathrm{cMpc/h})^3$ with 5040^3 dark matter particles of mass $\mathrm{M_p} = 1.061 \times 10^8 \mathrm{M}_\odot h^{-1}$ and Plank cosmology : $\mathrm{H_0} = 67.77$ km $\mathrm{s^{-1}Mpc^{-1}}$, $\Omega_\Lambda = 0.693$, $\Omega_M = 0.307$, $\sigma_8 = 0.8288$ (Planck Collaboration et al., 2016). The `P-Millennium` models the hierarchical growth of structures in the ΛCDM scenario in our model.

2. `GALFORM`, a semi-analytic model of galaxy formation and evolution. The `GALFORM` version used in this work is detailed in Lacey et al. (2016) and Baugh et al. (2019) and it is based on Cole et al. (2000).

   In short, initially `GALFORM` populates the dark matter halos extracted from a high redshift output of the `P-Millennium` with gas. Then, the gas is evolved tracking the merger histories of halos and several physical mechanisms are included, such as i) shock-heating and radiative cooling of gas inside halos; ii) formation of galactic disk with quiescent star formation ; iii) the triggering of starburst episodes due to disk instabilities and mergers in bulges; iv) AGN, supernovae and photoionization feedback to regulate the star formation rate; v) the chemical evolution of gas and stars. Additionally, the `GALFORM` version used in this work implements different stellar initial mass functions (IMFs) for quiescent star formation and starburst episodes (for further details in Lacey et al. (2016)). To ensure a proper resolution, in our model we consider only galaxies with stellar masses higher than $10^7 \mathrm{M}_\odot h^{-1}$, which roughly corresponds with dark halo masses around $10^{10} \mathrm{M}_\odot h^{-1}$ (~ 100 dark matter particles).

3. `FLaREON`[1] (Gurung-Lopez et al., 2018), an open source python code to predict Lyman α escape fractions and line profiles in minimal computational time. `FLaREON` is based on the radiative transfer Monte Carlo code `LyaRT`[2] (Orsi et al., 2012) that fully tracks the trajectory of Lyα photons in outflows with different gas geometries, hydrogen column densities ($\mathrm{N_H}$), macroscopic expansion velocities ($\mathrm{V_{exp}}$) and dust optical depths (τ_a). Briefly, `FLaREON` combines precomputed grids of $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$ and Lyα line profiles in the $\mathrm{N_H}$-$\mathrm{V_{exp}}$-τ_a space for several outflow geometries and different algorithms, such as multi-dimensional interpolation and machine learning, achieving high accuracy at a very low computational cost. We make use of `FLaREON` to include the RT inside galaxies in our model (see §4.2.1).

4. Radiative transfer of Lyα photons in the intergalactic medium. We estimate the IGM transmission for every galaxy depending on the local environment properties, such as the density, velocity and ionization state of the IGM (see §4.2.2).

In the following subsections we describe in detail the design of our LAE models.

---

[1] `https://github.com/sidgurun/FLaREON`

[2] `https://github.com/aaorsi/LyaRT`

4.2.1 Modeling the radiative transfer of Lyα photons inside galaxies.

The physics of Lyα photons escaping galaxies through galactic outflows are implemented with `FLaREON`. We focus on two outflow geometries: i) expanding homogenius Thin Shell (e.g. Verhamme et al., 2006; Orsi et al., 2012; Gurung López et al., 2018); ii) expanding Galactic Wind (Orsi et al., 2012; Gurung López et al., 2018) with a density gradient. Both geometries exhibit an empty cavity in the center of the geometry, where monochromatic Lyα photons are generated. We assume a constant temperature $T = 10^4 K$. These outflow geometries are detailed in Gurung-Lopez et al. (2018).

`GALFORM` galaxies are divided into two components, disk and bulge, with distinct galactic properties, such as metallicity, cold gas mass or star formation rate. Therefore, each component is assigned a unique $f_{\rm esc}^{\rm Ly\alpha}$ and Lyα line profile. This assumes that photons generated in a certain galaxy component only interact with that galaxy component.

We calculate outflow properties as in Gurung López et al. (2018). In particular, motivated by observational studies (e.g. Cazzoli et al., 2016), the outflow expansion velocity is computed as

$$\mathrm{V_{exp,c}} = \kappa_{V,c} \mathrm{SFR}_c \frac{r_c}{M_*}, \tag{4.1}$$

where the subscript c denotes the galaxy component (disk or bulge), $\mathrm{SFR_c}$ is the star formation rate in $\mathrm{M_\odot Gyr^{-1}} h^{-1}$ units, $\mathrm{r_c}$ is the half stellar mass radius in pMpc h^{-1}, $\mathrm{M_*}$ is the total stellar mas of the galaxy in $\mathrm{M_\odot} h^{-1}$ units. Additionally, $\kappa_{\rm V,c}$ are free dimensionless parameters regulating the efficiency of gas ejection.

In `FLaREON`, the Thin Shell and the Galactic Wind geometries present different density profiles, thus different column densities. The outflow neutral hydrogen column density of each component is computed as

$$\mathrm{N_{H,c}} = \begin{cases} \kappa_{N,c} \frac{\mathrm{M_{cold,c}}}{\mathrm{r_c^2}} & \text{Thin Shell} \\ \kappa_{N,c} \frac{\mathrm{M_{cold,c}}}{\mathrm{r_c V_{exp,c}}} & \text{Galactic Wind} \end{cases}, \tag{4.2}$$

where $\mathrm{M_{cold,c}}$ is the cold gas mass in $\mathrm{M_\odot} h^{-1}$ and $\kappa_{\rm N,c}$ are dimensionless free parameters. These are calibrated later in §4.3.

The free parameters presented above outline the outflow properties, regulating the Lyα luminosity distribution. In general, the parameters related to the quiescent star formation (galactic disks) give form to the faint end of the Lyα luminosity function. Meanwhile, the bright end is shaped by the free parameters regulating the starburst episodes (galactic bulges) (Lacey et al., 2016).

Finally, the dust absorption optical depth is simply computed as

$$\tau_{\mathrm{a,c}} = (1 - A_{\mathrm{Ly}\alpha}) \frac{E_{\odot}}{Z_{\odot}} \mathrm{N}_{\mathrm{H},c} Z_c, \tag{4.3}$$

where $\mathrm{E}_{\odot} = 1.77 \times 10^{-21}\mathrm{cm}^{-2}$ is the ratio $\tau_{\mathrm{a}}/\mathrm{N}_{\mathrm{H}}$ for solar metallicity, $A_{\mathrm{Ly}\alpha} = 0.39$ is the albedo at the Lyα wavelength, the solar metallicity is $\mathrm{Z}_{\odot} = 0.02$ (Granato et al., 2000) and Z_{c} is the metallicity of the cold gas in $\mathrm{Z}_{\odot}$ units.

4.2.2 Modeling the radiative transfer of Lyα photons inside the IGM.

While inside galaxies the losses of Lyα flux are due to dust absorption, in the IGM photons are scattered out of the line of sight by the neutral hydrogen. The total opacity of the IGM is given by (Dijkstra et al., 2007)

$$\tau_{IGM}(\lambda) = \sigma_0 \int_{R_{vir}}^{\infty} n_{\mathrm{H}}(s) \;\; \mathrm{X}_{\mathrm{H_I}}(s) \;\; \phi(\lambda, V_{\mathrm{shift}})\, ds, \tag{4.4}$$

where λ is the wavelength, s is the proper distance to the galaxy where the photon is emitted, n_{H} is the IGM hydrogen number density, $\mathrm{X}_{\mathrm{H_I}}$ is the fraction of neutral hydrogen in the IGM and $\phi(\lambda, V_{\mathrm{shift}})$ is the Voigt profile Doppler-shifted by the velocity between the emitting galaxy and the IGM.

We compute the IGM transmission of each galaxy, which depends on the local environment. To do so we compute $n_{\mathrm{H}}(\overrightarrow{x})$, $\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x})$ and the hydrogen velocity field, $V_{\mathrm{H}}(\overrightarrow{x})$ from our simulation. Due to disk storage limitations, the dark matter particles of `P-Millennium` at the snapshots used in this work were not saved. Hence, these quantities are computed from the halo catalogs as shown in the next subsections.

Hydrogen number density field.

We assume that density of hydrogen, $\rho_{\mathrm{H}}(\overrightarrow{x})$, is coupled to the dark matter density, $\rho_{\mathrm{DM}}(\overrightarrow{x})$. Thus,

$$n_{\mathrm{H}}(\overrightarrow{x}) = x_{\mathrm{H}} \frac{\Omega_{\mathrm{b}}}{\Omega_{\mathrm{DM}}} \rho_{\mathrm{DM}}(\overrightarrow{x}) / m_{\mathrm{H}}, \tag{4.5}$$

where $x_{\mathrm{H}} = 0.74$ is the hydrogen fraction of baryonic matter in the Universe, Ω_{b} and Ω_{DM} are the densities of baryons and dark matter respectively and m_{H} is the hydrogen mass.

In order to compute $\rho_{\mathrm{DM}}(\overrightarrow{x})$, it is useful to define the overdensity field of a given quantity, $\delta_a(\overrightarrow{x})$, as

$$\delta_a(\overrightarrow{x}) = \frac{\rho_a(\overrightarrow{x}) - \langle \rho_a \rangle}{\langle \rho_a \rangle}, \tag{4.6}$$

where $\rho_a(\overrightarrow{x})$ is the density field of some quantity and $\langle \rho_a \rangle$ is its average. Additionally, the definition of bias for a given dark matter halo mass, $b(M)$, is

$$\delta_{\mathrm{halos}}(\overrightarrow{x}) = b(M) \delta_{\mathrm{DM}}(\overrightarrow{x}), \tag{4.7}$$

where $\delta_{\mathrm{halos}}(\overrightarrow{x})$ is the overdensity field of the dark matter halos and $\delta_{\mathrm{DM}}(\overrightarrow{x})$ is the dark matter overdensity field.

By combining eq. 4.6 and eq. 4.7, $\rho_{\mathrm{DM}}(\overrightarrow{x})$ can be expressed as

$$\rho_{\mathrm{DM}}(\overrightarrow{x}) = \left(\frac{\delta_{\mathrm{halos}}(\overrightarrow{x})}{b_{\mathrm{eff}}(\overrightarrow{x})} + 1 \right) \langle \rho_{\mathrm{DM}} \rangle \ , \tag{4.8}$$

where we have defined the effective bias $b_{\mathrm{eff}}(\overrightarrow{x})$ in each cell as

$$b_{\mathrm{eff}}(\overrightarrow{x}) = \frac{\int b(M) \frac{dN}{dM}(\overrightarrow{x})\, dM}{\int \frac{dN}{dM}(\overrightarrow{x})\, dM}, \tag{4.9}$$

where $\frac{dN}{dM}(\overrightarrow{x})$ is the halo mass distribution in each cell.

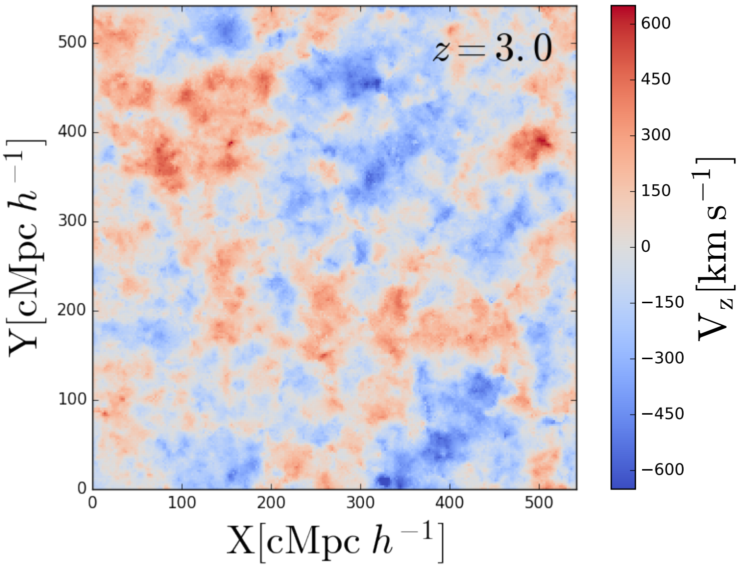


Fig. 4.2 Velocity along the line of sight (Z axis) spatial distribution at redshift $z = 3$ in a slice of $\sim 2\mathrm{cMpc}h^{-1}$ width. Velocities pointing to the observer are shown in blue, while regions distancing from the observer are drawn in red.

Velocity field.

We assume that the motion of the dark matter and gas are the same, i.e., $V_{\mathrm{H}}(\overrightarrow{x}) = V_{\mathrm{DM}}(\overrightarrow{x})$. In practice, we divide our simulation box into smaller volumes and compute $V_{\mathrm{DM}}(\overrightarrow{x})$ from the halo catalog simply as the median of the velocity distribution of halos within the subvolume.

In Fig.4.2 we show the matter velocity along the line of sight (chosen arbitrarily as the Z coordinate of the simulation box) at redshift 3.0 in a slice of width $2\mathrm{cMpc}h^{-1}$ of our simulation box. The volume is divided in big chunks with coherent positive motion along the line of sight (red) and negative motion (blue). The typical scales of these areas are hundreds of comoving megaparsecs. Meanwhile, the transition between these regions is relatively small. This causes great contrasts of velocity along the line of sight on scales below $\sim 10\mathrm{cMpc}h^{-1}$. We have checked that this behaviour is also found at the other redshifts studied in this work.

Fraction of neutral hydrogen field.

Star forming galaxies and active galactic nuclei (AGN) are the main sources of ionizing photons (photons wavelength shorter than 912Å) in the Universe (Kimm and Cen,

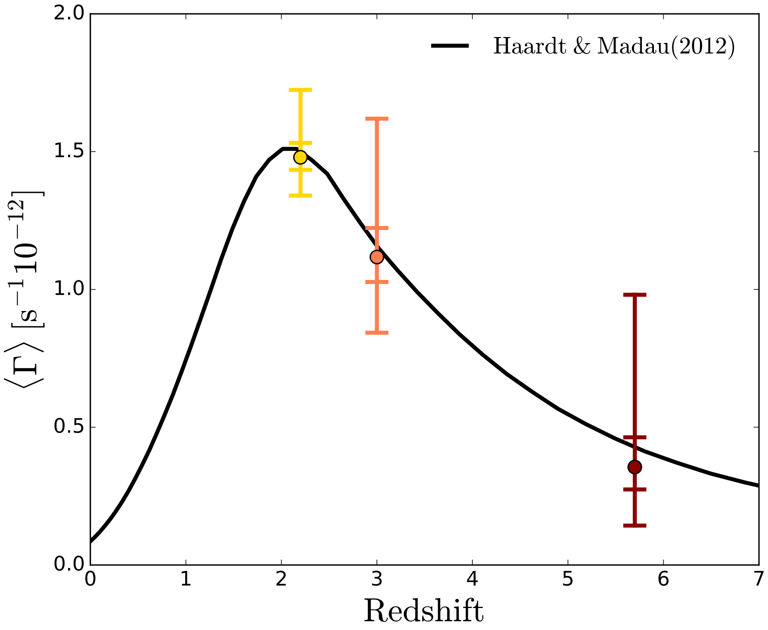


Fig. 4.3 Cosmic photoionization rate (Γ) as a function of redshift. The solid black curve shows the Haardt and Madau (2012), while the dots show the median value of Γ in our simulation at different redshifts. The error bars correspond to the percentiles 5, 32, 68 and 95 of Γ from the lowest to the highest.

2014). Thus the photoionization field, $\Gamma(\vec{x})$, is coupled to the number density field of these objects. In principle, regions with a high rate of ionizing photons, i.e. close to these sources, will be more ionized, thus lowering the IGM opacity to Lyα photons.

In order to compute $\Gamma(\vec{x})$ we assume that the ionizing radiation is produced only by galaxies. This assumption works well at high redshifts since previous studies have shown that other sources of ionizing photons, such as AGNs or QSOs, are not sufficiently abundant to significantly contribute to $\Gamma(\vec{x})$ (Parsa et al., 2018). We do not expect this to significantly affect our result since, as described later, $\Gamma(\vec{x})$ is calibrated to reproduce the observed mean Γ at the different epochs studied here.

We compute $\Gamma(\vec{x})$ as the superposition of every ionizing field generated by each galaxy, i.e,

$$\Gamma(\vec{x}) = \sum_i \frac{\sigma_0 \dot{Q}_{\mathrm{H},i} f_{\mathrm{esc}}^{\mathrm{ion}}}{4\pi|\vec{x} - \vec{x}_i|^2} \frac{\beta}{\beta - 3} G(|\vec{x} - \vec{x}_i|), \qquad (4.10)$$

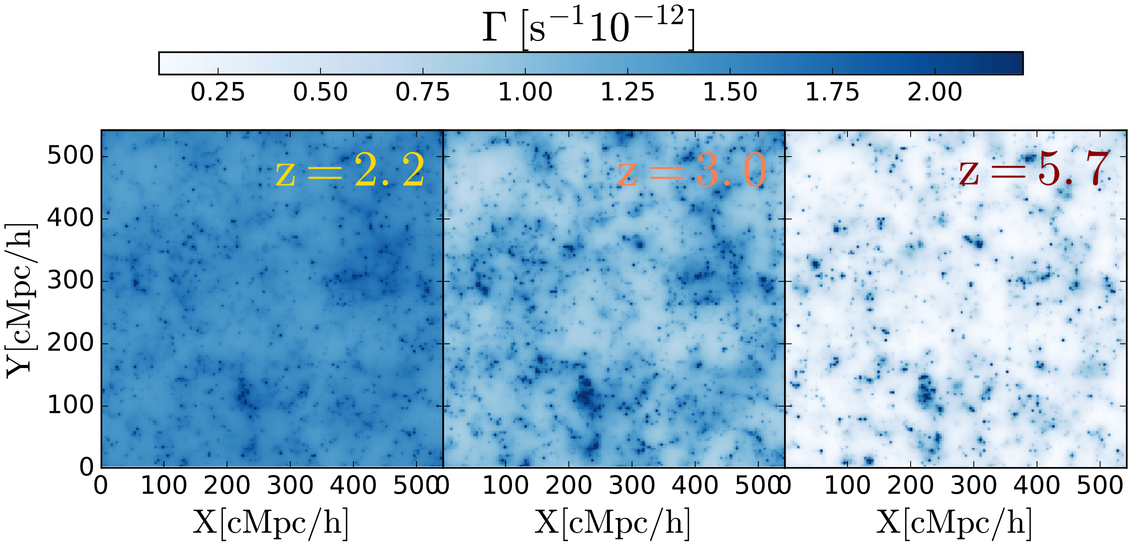


Fig. 4.4 Spatial distribution of the photoionization rate in different snapshots of the simulation (as labeled) in slices of $\sim 2\mathrm{cMpc}h^{-1}$ width.

where the sum is over all the galaxies in the box, $\overrightarrow{x}_i$ is the location of each galaxy, $\overrightarrow{x}$ is the position where ionizing field in evaluated. Additionally, we assume a global escape fraction of ionizing photons $f_{\mathrm{esc}}^{\mathrm{ion}} = 0.1$ (Kimm and Cen, 2014) and $\dot{Q}_{\mathrm{H},i}$ is the total luminosity of ionizing photons given by `GALFORM` for each galaxy. We assume that the SED of the galaxy takes the form $J(\nu) = \nu^{\beta}$ in bluer parts of the hydrogen ionization frequency threshold, ν_H. We also assume that the photoionization cross section as $\sigma_{\mathrm{H}}(\nu) = \sigma_0(\nu/\nu_{\mathrm{H}})^{-3}$ with $\sigma_0 = 6.3 \times 10^{-18} cm^{-2}$. Finally, the function $G(|\overrightarrow{x} - \overrightarrow{x}_i|)$ takes into account the fact that photons emitted by a single galaxy do not reach every point in space and is given by

$$G(|\overrightarrow{x} - \overrightarrow{x}_i|) = \begin{cases} 0 & \text{if } |\overrightarrow{x} - \overrightarrow{x}_i| > R_{ion,i} \\ 1 & \text{if } |\overrightarrow{x} - \overrightarrow{x}_i| < R_{ion,i} \end{cases}, \tag{4.11}$$

where $R_{ion,i}$ is the radius of the sphere centered in the location of the galaxy i. We use a similar expression to the Strömgren radius to compute $R_{ion,i}$:

$$R_{ion,i} = K_a(\dot{Q}_{H,i}/10^{55}s^{-1})^{1/3}, \tag{4.12}$$

where K_a is a free parameter. Increasing (decreasing) K_a leads to greater (lower) $R_{ion,i}$. Hence, at each point, more (less) galaxies contribute to the the ionizing radiation field,

Table 4.1 Values of the best fitting K_a used to compute the ionization field in our model.

redshift	$\log K_a$[cMpc h^{-1}]
2.2	2.79
3.0	3.65
5.7	4.48

which augments (lowers) the $\langle\Gamma\rangle$ of the simulation box.

We determine the K_a values at each redshift by fitting the $\langle\Gamma\rangle$ of our simulations to the $\langle\Gamma\rangle$ given by Haardt and Madau (2012). The best fitting values are listed in Table 4.1. We find that K_a decreases with redshift. Since the number density of ionizing sources decreases from $z = 2.2$ to 5.7, the volume ionized by a single galaxy must be increased with redshift in order to fit observations.

Fig.4.3 displays a comparison between the observed $\langle\Gamma\rangle$ and the $\Gamma(\overrightarrow{x})$ distribution (percentiles 5, 32, 50, 68 and 95) of our model.

Fig.4.4 shows the spatial variations of $\Gamma(\overrightarrow{x})$ over cosmic time in the same slice of width ~ 2cMpch^{-1}. By construction, $\Gamma(\overrightarrow{x})$ reproduces observations quite well. At $z = 2.2$ $\langle\Gamma\rangle$ peaks and decreases towards higher redshifts. The dispersion of $\Gamma(\overrightarrow{x})$ evolves with redshift too. While at high redshift the photoionization field exhibits a complex structure with high contrasts, as the Universe evolves, it becomes smoother.

Finally, we compute the neutral hydrogen fraction field as in Dijkstra and Loeb (2009) :

$$\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x}) = 1 + a(\overrightarrow{x})/2 - \left[a + (a(\overrightarrow{x})/2)^2\right]^{1/2}, \tag{4.13}$$

with $a(\overrightarrow{x}) = \Gamma(\overrightarrow{x})/n_{\mathrm{H}}(\overrightarrow{x})\alpha_{\mathrm{rec}}$ where the case-A[3] recombination coefficient is given by $\alpha_{\mathrm{rec}} = 4.2 \times 10^{-13}(T_{\mathrm{gas}}/10^4K)^{0.7}cm^3s^{-1}$ and $T_{\mathrm{gas}} = 10^4K$ is assumed.

---

[3] Case-A recombination assumes that the medium is optically thin to ionizing photons.

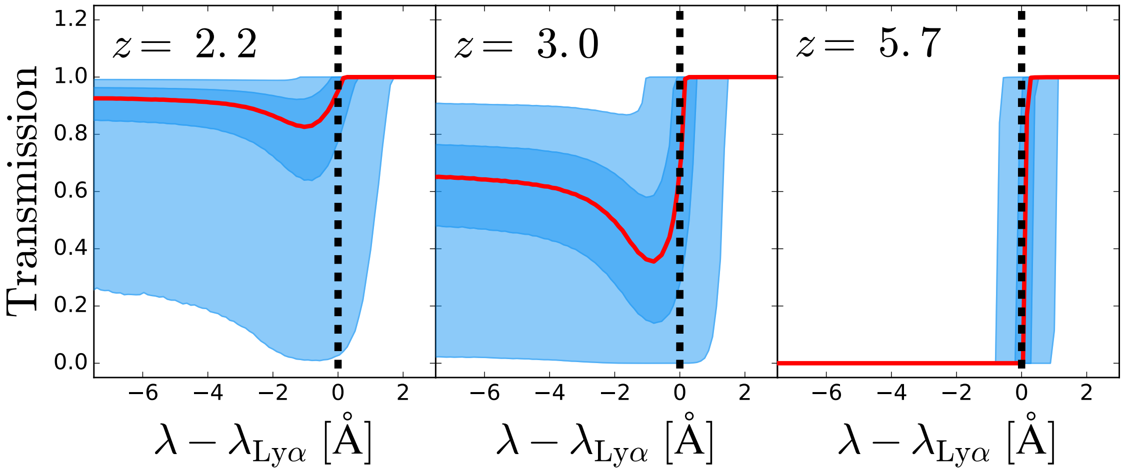


Fig. 4.5 Average IGM transmission around Lyα for different redshift bins (2.2, 3.0 and 5.7 from left to right). In red we show the median IGM transmission, while in dark blue and light blue we show the 1σ and 2σ ranges of the distribution.

IGM transmission.

We compute $n_H(\overrightarrow{x})$, $V_H(\overrightarrow{x})$ and $\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x})$ on a grid of 250^3 cells with side $\sim$ 2cMpc h^{-1}. Then, by linear interpolation, we reevaluate these fields on a grid with higher spatial resolution along an arbitrary direction chosen as the line of sight. In particular, the new grid is composed of 250×250×3000 cells, where the size of the cells along the line of sight is $\mathrm{L_{cells}} \sim$ 0.2cMpc/h. We choose this grid size to ensure good signal to noise in the computed field and the IGM transmission curves with high enough resolution in frequency space.

Following eq. 4.4, the IGM optical depth in each cell in the simulation box rest frame is given by

$$\tau_{\mathrm{IGM}}(\lambda) = \sum_{l=l_{\mathrm{gal}}} \sigma_0 L_{\mathrm{cells}} n_{\mathrm{H_I}}(i,j,l)\ \mathrm{X}_{\mathrm{H_I}}(i,j,l)\ \phi(\lambda, V_{\mathrm{shift}}(i,j,l)), \tag{4.14}$$

where the cell indices $\{i,\ j\}$ and $\{l\}$ are perpendicular and parallel, respectively, to the line of sight. Moreover, l starts at the cell where the galaxy lies, l_{gal}, and iterates towards the observer along the line of sight direction. Additionally, V_{shift} is the relative velocity along the line of sight between the cell containing the galaxy (i,j,$l = l_{\mathrm{gal}}$) and the IGM in the iterated cell (i,j,l), given by

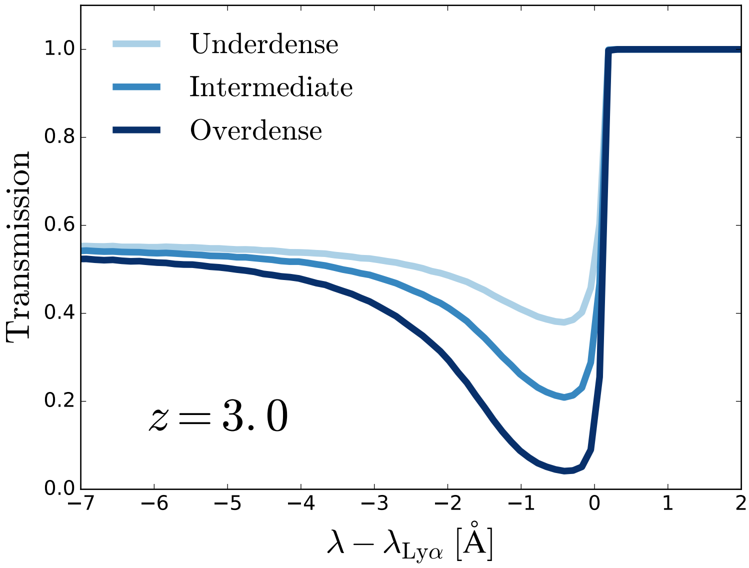


Fig. 4.6 Median IGM transmission around the Lyα wavelength for LAE samples hosted regions with different density. LAE living at the 1/3 of the lowest densities are shown in light blue, while in blue from the 1/3 to the 2/3 and dark blue for the LAE lying in the 1/3 densest environments.

$$V_{\mathrm{shift}}(i,j,l) = V_{\mathrm{H}}(i,j,l) + H_{\mathrm{Hubble}}(z) \times (l_{\mathrm{gal}} - l) \times L_{\mathrm{cells}}, \tag{4.15}$$

where $H_{\mathrm{Hubble}}(z)$ is the Hubble constant at redshift z.

Finally, the IGM transmission at each position is computed as $T(\lambda) = e^{-\tau_{\mathrm{IGM}}(\lambda)}$. In Fig. 4.5 we show the median and dispersion of the IGM transmission at different redshifts for a sample of the 150000 `GALFORM` galaxies with the highest specific star formation rate (sSFR) as representative of an emission line galaxy population. In general, we find a good qualitative agreement between the shape of our mean transmission curves and the mean transmission curves obtained by Laursen et al. (2011a) computed from hydrodynamical simulation with RT physics implemented. The IGM absorbs photons bluer than $\lesssim 1216$Å. Moreover, as galaxies lie in overdense regions, the IGM opacity is higher close to the galaxy, causing the drop in the transmission close to Lyα wavelength. Then the IGM transmission flattens to the IGM cosmic transmission.

Additionally, the optical depth of the IGM evolves with redshift, producing higher transmissions at lower redshifts. This becomes dramatic at $z = 5.7$, where the IGM transmission goes below 1% at bluer frequencies than Lyα.

The large dispersion of the IGM transmission reflects the complex variety of environments surrounding galaxies. In order to test this idea, we rank our LAE samples in IGM density $\rho_{\rm H}$ and split them into 3 subsamples: underdense (below the percentile 33rd of density), intermediate (between the 33 and 66 percentiles) and overdense (above the 66 percentile). In Fig. 4.6 we show the median IGM transmission at $z = 3.0$ for these subsamples. In general the behaviour of the three population is the same. At redder wavelengths than Lyα the transmission is 1. Meanwhile, at bluer wavelengths far from Lyα ($\lambda \sim 1210$Å) the transmission converges to the mean IGM transmission. However, at blue wavelengths around Lyα ($\lambda \sim 1214$Å) the density has a great impact on the IGM transmission. We find that LAEs hosted in denser environments exhibit lower IGM transmission than their counterparts in low density regions. The transmission is ~ 0.4, ~ 0.2 and ~ 0.1 for the underdense, intermediate and overdense environments. For completeness, we also did this analysis at the other snapshots of our simulations, finding the same trend. The typical transmission at $z = 2.2$ around Lyα is ~ 0.9, ~ 0.85 and ~ 0.8 for the underdense, intermediate and overdense environments. Meanwhile, at $z = 5.7$ the transmission remains below 1% even in the underdense regions.

4.2.3 The observed Lyα luminosity.

The observed Lyα luminosity is a convolution of galactic and IGM properties. We compute the observed Lyα luminosity as follows.

Physical processes taking place inside galaxies are implemented as in (Gurung López et al., 2018). For each galactic component: i) the intrinsic Lyα luminosity, $\rm L^{0,c}_{Ly\alpha}$ (where the superindex "c" denotes the disk or bulge component), is predicted by `GALFORM` from the instantaneous star formation rate; ii) the outflow properties are computed with equations 4.1, 4.2 and 4.3; iii) the outflow Lyα escape fraction $f^{c}_{\rm esc}$ and line profile $\Phi^{\rm c}(\lambda)$ are predicted with `FLaREON`; iv) the Lyα luminosity escaping the ISM is computed as

$$\rm L^{c}_{Ly\alpha} = L^{0,c}_{Ly\alpha} \mathit{f}^{c}_{esc}. \tag{4.16}$$

The IGM transmission is calculated in the simulation rest frame, as discussed in the previous section. However, the galaxy rest frame mismatches, in general, the simulation

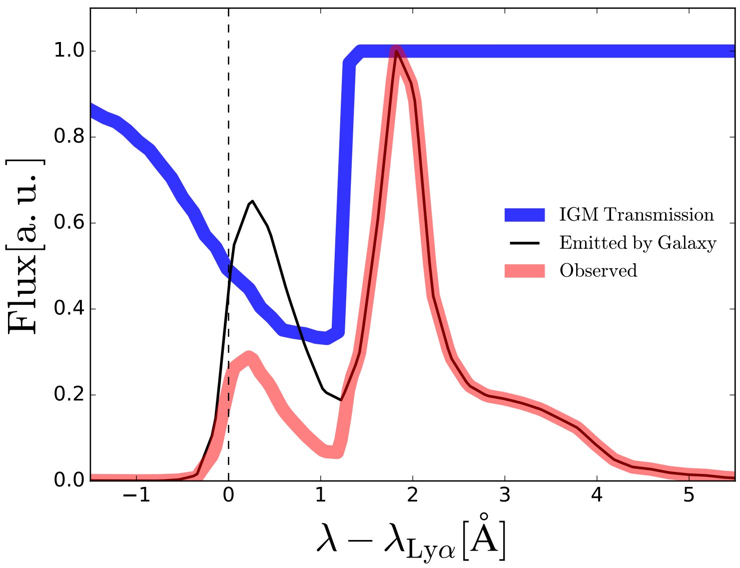


Fig. 4.7 Example of the interaction between the IGM and a galaxy approaching the IGM along the line of sight. We show in blue the IGM transmission at the galaxy position, in solid black the Lyα line profile emerging from the galaxy after the interaction with the ISM and in thick red the observed Lyα line profile after the IGM absorption.

rest frame due to galaxy peculiar velocities. Hence, the IGM transmission for a given galaxy is Doppler shifted by the peculiar velocity of the galaxy along the line of sight. Then, for each component, the fraction of photons that travel unscattered through the IGM, $f_{esc}^{\mathrm{IGM,c}}$, is computed as

$$f_{\mathrm{esc}}^{\mathrm{IGM,c}} = \frac{\int \Phi^{\mathrm{c}}(\lambda) \times \mathrm{T}^{\mathrm{gal}}(\lambda) d\lambda}{\int \Phi^{\mathrm{c}}(\lambda) d\lambda}, \tag{4.17}$$

where $\mathrm{T}^{\mathrm{gal}}(\lambda)$ is the IGM transmission at the galaxy position and rest frame and Φ^{c} is the Lyα line profile of each galaxy component.

In Fig. 4.7 we illustrate how $f_{\mathrm{esc}}^{\mathrm{IGM,c}}$ is computed. In this particular case the IGM and the galaxy are approaching along the line of sight. As shown previously, in the rest frame, the IGM absorbs photons bluer than Lyα. However, due to the relative motion between the galaxy and the IGM along the line of sight, the IGM transmission (blue curve) is redshifted in the galaxy rest frame. Hence, the Lyα photons emerging from the ISM (black curve) are partially absorbed. As a result the observed Lyα line profile (red curve) is modified. $f_{\mathrm{esc}}^{\mathrm{IGM,c}}$ is calculated as the ratio between the integrals of the observed and the emerging line profiles.

Table 4.2 Free parameters as defined in equations 4.1 and 4.2 after the calibration with the observed luminosity function for different geometries and redshifts.

redshift	Geometry	$\log \kappa_{V,disk}$	$\log \kappa_{V,bulge}$	$\log \kappa_{N,disk}$	$\log \kappa_{N,bulge}$
$z = 2.2$	Thin Shell	4.43	4.27	-12.33	-12.11
	Galactic Wind	4.14	4.80	-8.01	-5.58
$z = 3.0$	Thin Shell	3.78	4.23	-12.73	-12.06
	Galactic Wind	4.04	4.47	-8.10	-5.69
$z = 5.7$	Thin Shell	4.55	3.48	-14.03	-12.24
	Galactic Wind	4.35	3.28	-9.62	-6.85

Then, for each galaxy, the observed Lyα luminosity is computed as

$$\mathrm{L}_{\mathrm{Ly}\alpha} = \mathrm{L}^{\mathrm{Disk}}_{\mathrm{Ly}\alpha} \times f^{\mathrm{IGM,Disk}}_{\mathrm{esc}} + \mathrm{L}^{\mathrm{Bulge}}_{\mathrm{Ly}\alpha} \times f^{\mathrm{IGM,Bulge}}_{\mathrm{esc}}. \tag{4.18}$$

Moreover, we estimate the rest frame Lyα equivalent width, EW, as

$$\mathrm{EW} = \mathrm{L}_{\mathrm{Ly}\alpha}/\mathrm{L}_{\mathrm{continuum}}, \tag{4.19}$$

where $\mathrm{L}_{\mathrm{continuum}}$ is the continuum luminosity per unit of wavelength around Lyα computed by `GALFORM`. This quantity is based on the evolution of the composite stellar population of each modeled galaxy. From now on, unless stated otherwise, we define LAEs in our model as galaxies with EW > 20Å.

4.3 Lyman-α emitters

4.3.1 Model calibration.

In these section we briefly discuss how the free parameters described in Eq.4.1 and 4.2 are calibrated. For further details we refer the reader to Gurung López et al. (2018).

In short, for each redshift and outflow geometry, we perform an MCMC analysis using the open source `Python` library `emcee` (Foreman-Mackey et al., 2013a) to estimate the values of $\kappa_{\mathrm{N,c}}$ and $\kappa_{V,c}$ that best reproduce the observed LAE luminosity function (LF) at different redshifts. In particular, at $z = 2.2$ we fit the observed LF from Cassata et al. (2011), Konno et al. (2016) and Sobral et al. (2017a), while at $z = 3.0$ we fit the LFs from Ouchi et al. (2008) and Cassata et al. (2011). Finally, at $z = 5.7$ we model the LFs of Ouchi et al. (2008) and Konno et al. (2018). The free parameter values

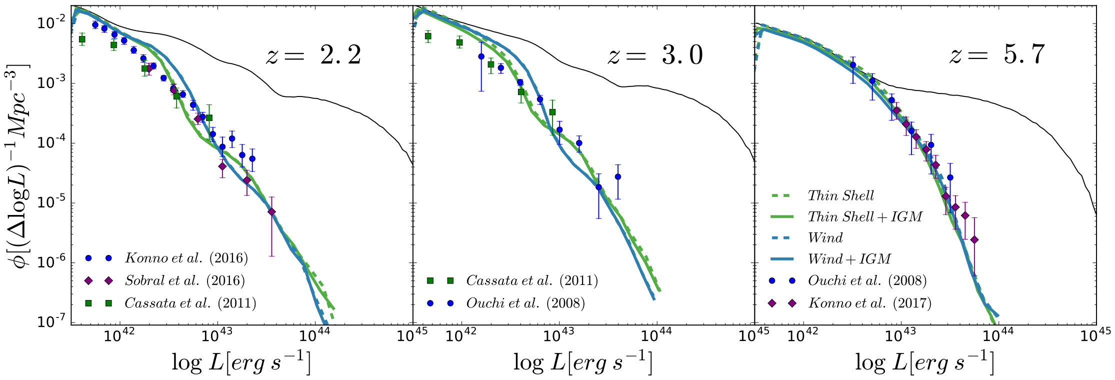


Fig. 4.8 Comparison between our model's LAE luminosity function (LF) and observations at different redshifts. The intrinsic Lyα LF is shown in thin black, the Thin Shell and Galactic Wind geometries are shown in green and blue dashed lines respectively. The colored dashed lines correspond to models with RT in the ISM but without RT in the IGM, while the colored solid lines include the RT in the ISM ad IGM. Different observational data sets are shown in symbols according to the legends.

are listed in Table 4.2. Additionally, the resulting $\mathrm{N_H}$, $\mathrm{V_{exp}}$ and τ_a distributions are discussed in Appendix 4.9.1.

In Fig. 4.8 we compare the observed LF at several redshifts with our models. For illustration we show the `GALFORM` intrinsic LAE LF (black thin line), which over predicts the number of LAEs over the full luminosity range at $z = 2.2$ and $z = 3.0$. Hence, the total Lyα escape fraction $f_{\rm esc}^{\rm Ly\alpha}$ (galaxy+IGM) must be < 1. However at redshift $z = 5.7$, while bright intrinsic LAE surpass observations, the intrinsic number counts of faint LAEs resembles observations, implying $f_{\rm esc}^{\rm Ly\alpha} < 1$ and $f_{\rm esc}^{\rm Ly\alpha} \sim 1$ respectively.

In general, after calibration, our models (colored solid lines) match the observed LAE LF at all redshifts by construction. The good agreement is remarkable at $z = 5.7$. The Thin Shell geometry matches slightly better observations than the Galactic Wind. Additionally, we show the LAE LF of our calibrated model excluding the IGM absorption (colored dashed lines). Although the LAE number counts of the model without IGM exceeds the abundance of LAE in the complete models at every Lyα luminosity, their LAE LF are very similar. This points to the fact that the RT inside galaxies is the main driver shaping the observed LAE LF.

4.3.2 LAE samples

Throughout this work we analyze and compare the properties of different LAE samples to highlight the RT selection effects. Here we describe how the samples are built.

Full Lyα emitters samples.

Full Lyα emitters (FLAE). These samples represent the observed LAE population. They include all the radiative transfer processes explained above (ISM + IGM). The FLAE samples are derived from the full `GALFORM` population. The outflow properties of the `GALFORM` galaxies are computed with the calibrated free parameters (listed in Table 4.2). Then we assign to each galaxy a Lyα luminosity (as described in §4.2.3). Note that these samples, by construction, reproduce the observed LF. We rank these populations by Lyα luminosity and perform a number density cut of $4 \times 10^{-3}(\mathrm{cMpc}\ h^{-1})^{-3}$. The chosen number density cut is arbitrary. We obtain similar results for higher and lower number density cuts.

Partial Lyα emitters samples.

Partial Lyα emitters (PLAE). These samples include galactic RT physics but lack the IGM absorption. They are also subsamples of the full `GALFORM` galaxy population. PLAEs can be seen as the LAE population that would be observed if the IGM was completely transparent. In particular, $\mathrm{N_H}$ and $\mathrm{V_{exp}}$ are computed with the same calibration as the models with full RT (FLAE samples). However, in contrast to FLAE populations where the Lyα luminosity is computed through Eq. 4.18, in the PLAE samples, the observed $\mathrm{L_{Ly\alpha}}$ is computed simply as

$$\mathrm{L_{Ly\alpha}} = \mathrm{L^{Disk}_{Ly\alpha}} + \mathrm{L^{Bulge}_{Ly\alpha}}, \tag{4.20}$$

where the ISM RT is included in $\mathrm{L^{Disk}_{Ly\alpha}}$ and $\mathrm{L^{Bulge}_{Ly\alpha}}$.

Finally, we rank galaxies by $\mathrm{L_{Ly\alpha}}$ and make the same number density cut as in the FLAE samples. Note that, by construction, the intrinsic galaxy population and properties are identical for FLAE and PLAE samples. However, the FLAE samples include the IGM selection effects. Therefore, the comparison between these samples sheds light on the impact of the IGM.

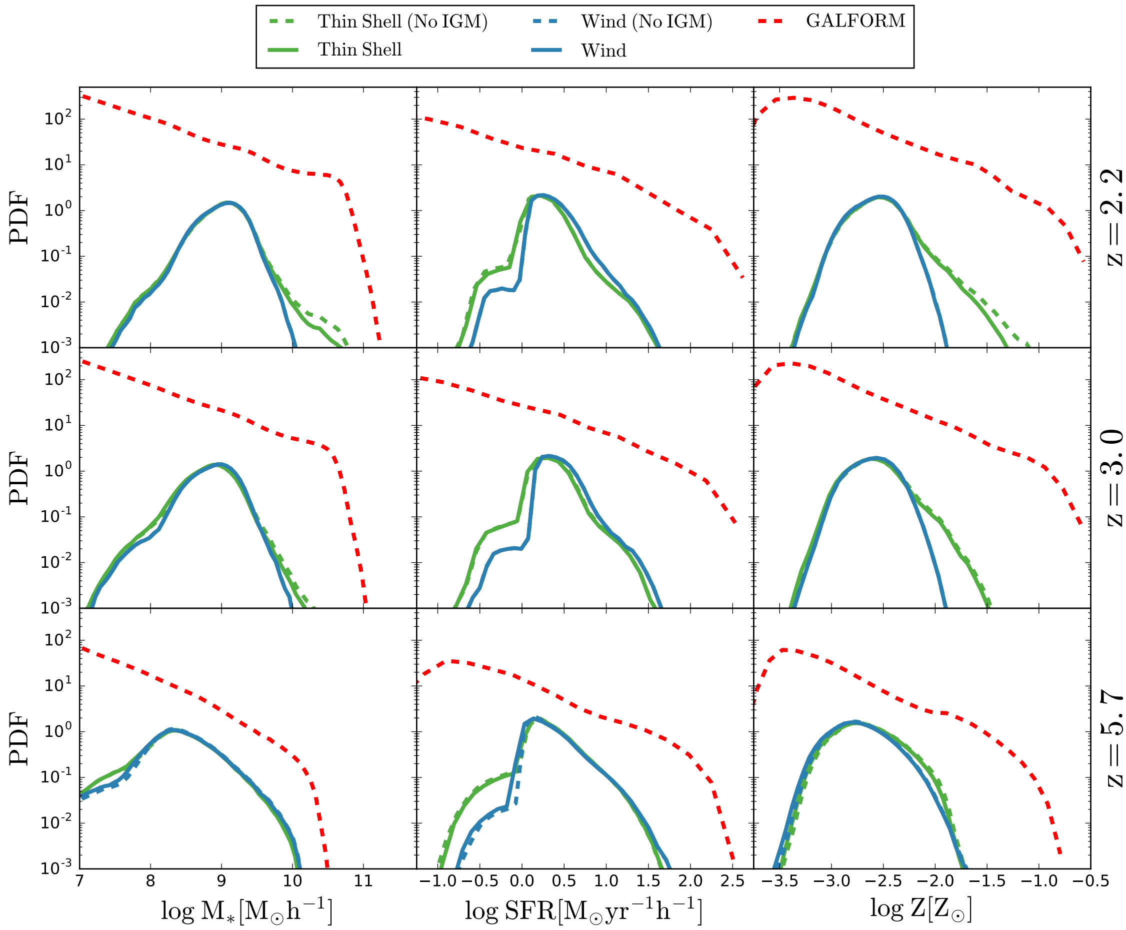


Fig. 4.9 Comparison between the galaxy properties distribution (stellar mass, star formation rate and metallicity from left to right) of our different galaxy samples at several redshifts (2.2, 3.0, 5.7 from top to bottom). In dashed red we show the galaxy property distributions of the full `GALFORM` catalog. In dashed and solid green we display the galaxy properties distribution of the Thin Shell excluding and including the RT in the IGM. Meanwhile, blue shows the same as green but implementing the Galactic Wind geometry.

Shuffled Lyα emitters samples.

Shuffled Lyα emitters (SLAE) samples are built to exhibit the same clustering as the FLAE samples if no LAE-IGM coupling is found. Throughout this work we analyze the clustering and the galactic properties of LAEs and how the IGM affects them. Since the FLAE and PLAE are computed from the same galaxy population they are useful to understand how the IGM shapes the galactic properties. As seen later on, the FLAE and PLAE exhibit different mass functions. This causes that each population displays a different clustering, and in particular, different bias. Additionally, if the large scale IGM properties are coupled to the Lyα observability, further clustering distortions are expected (Zheng et al., 2011). Therefore, in order to study the LAE-IGM coupling it would be desirable to use samples with the same bias.

SLAE samples are derived from the `P-Millennium` halo catalog mimicking the FLAE halo mass functions. In detail:

1. We separate our full RT LAE population into centrals and satellites. We find that at all redshifts and gas geometries, central galaxies constitute $\sim 98\%$ of the full RT LAE samples.
2. Comparing with the dark matter halo catalogs of the simulation we compute the fraction of halos occupied by LAEs as a function of mass (HOD) individually for the central and satellite population.
3. From the halo catalog we select central halos reproducing the HOD of the central LAE sample using a uniform random distribution.
4. We determine the number of LAE satellites hosted in each dark matter halo with a random Poisson distribution with mean equal to the satellite's HOD evaluated at the mass of the halo. Following Jímenez et al. (in prep.), we assign to the satellite the same location as the central halo and restrict our clustering analysis to the 2-halo-term scales.
5. We combine the new satellite and central population in a single SLAE sample.

By construction, SLAE are free of the IGM selection effects and exhibit the same halo mass distribution as their full RT progenitors. Therefore, if the IGM is not shaping the LAE spatial distribution, the clustering in the SLAEs and full RT LAE samples should be the same[4] on scales greater than the 1-halo term.

---

[4] Note that, by construction, the SLAE samples do not isolate the IGM-LAE coupling, but it also includes the effects of assembly bias (Contreras et al., 2019). This might cause differences between the clustering of the FLAE and SLAE populations. To check the assembly bias impact on the clustering of our LAE populations we built SLAE samples from the PLAE samples. No significant difference was found between the clustering of these populations, i.e., there was no assembly bias evidence. Therefore, we assume that the difference between the FLAE and SLAE (computed from the FLAE) samples are due to the IGM-LAE coupling.

4.4 LAE galaxy properties.

In this section we briefly study the selection function of LAEs. First, we analyze which galaxies would be observed as LAEs if the IGM was completely transparent to photons around Lyα. For this goal, we contrast the full `GALFORM` galaxy population with the PLAE samples, which includes RT only in the ISM. In the second case, we characterize the IGM impact by directly comparing the PLAE and the full RT LAE sample. The variations among these samples are caused by the IGM, since the only difference between the PLAE and FLAE samples is that FLAE also include RT in the IGM.

In Fig.4.9 we compare the galaxy property distributions between our Lyα flux selected samples and the full galaxy population predicted by `GALFORM` at different redshifts. We define a Lyα-weighted average gas metallicity for a galaxy as

$$\mathrm{Z} = \frac{\mathrm{Z}^{\mathrm{Disk}}\mathrm{L}^{\mathrm{Disk}}_{\mathrm{Ly}\alpha} + \mathrm{Z}^{\mathrm{Bulge}}\mathrm{L}^{\mathrm{Bulge}}_{\mathrm{Ly}\alpha}}{\mathrm{L}^{\mathrm{Disk}}_{\mathrm{Ly}\alpha} + \mathrm{L}^{\mathrm{Bulge}}_{\mathrm{Ly}\alpha}}. \tag{4.21}$$

We find very little difference between the FLAE and PLAE populations. This indicates that the IGM does not induce significant selection effects on the galaxy properties of LAEs. Moreover, independently of the outflow geometry and cosmic time, galaxies with strong Lyα emission present moderate stellar mass and SFR and low metallicity. We find that models implementing different outflow geometries behave in a similar fashion. Still there are tiny differences between them.

The typical stellar mass of galaxies observed as LAEs also evolve over cosmic history. In general, we find the same trend in the LAEs samples as in the full galaxy population: LAEs at lower redshift exhibit higher stellar content than their homologous at higher redshift. In detail, the M_* distributions peak around moderate masses; $10^{8.5}\mathrm{M}_\odot h^{-1}$, $10^{9}\mathrm{M}_\odot h^{-1}$ and $10^{9.2}\mathrm{M}_\odot h^{-1}$ at redshift 5.7, 3.0 and 2.2 respectively. The shapes of the distributions are very similar between $z = 2.2$ and $z = 3.0$. The dynamical range of M_* at these redshifts is very similar to the observed LAEs (Oyarzún et al., 2017). Additionally, very small differences can be found between the two outflow geometries implemented in this work. At $z = 2.2$ and 3.0 the Thin Shell predicts a greater abundance of massive galaxies in comparison with the Galactic Wind. This trend disappears at $z = 5.7$, when both stellar mass distribution are almost identical.

Meanwhile, the SFR of the full galaxy population also evolves with redshift. `GALFORM` predicts a progressive increase in the SFR from redshift 5.7 to 2.2 (middle column of

Fig.4.9). However, LAE populations exhibit a moderate SFR distribution that remains almost frozen through cosmic time. For all redshifts the SFR peaks at $10^{0.25}\mathrm{M}_{\odot}h^{-1}yr^{-1}$, well below the galaxies with the highest SFR ($\sim 10^{2.5}\mathrm{M}_{\odot}h^{-1}yr^{-1}$). We find no significant evolution from $z = 2.2$ to 3.0, while at $z = 5.7$ the distribution becomes a little bit broader.

The outflow geometry has a significant impact on the SFR distribution of LAEs. In particular, at all redshifts, the LAE samples characterized by the Thin Shell exhibit lower SFR than those using the Galactic Wind. This is caused by several reasons. First, the recipes to link galaxy proprieties to outflow properties are different between both geometries (see Eq.4.1 and Eq.4.2). Second, the escape fraction of Lyα photons depends strongly on the geometry (Gurung López et al., 2018; Gurung-Lopez et al., 2018).

`GALFORM` also predicts an evolution through cosmic time of the metal abundance in galaxies (right column of Fig.4.9). In fact, galaxies at lower redshifts exhibit higher metallicity. This is a consequence of the consecutive events of SFR that pollute the initially pristine interstellar medium. The typical maximum metallicities are $\mathrm{Z} \sim 10^{-0.75}\mathrm{Z}_{\odot}$, $\sim 10^{-0.6}\mathrm{Z}_{\odot}$ and $\sim 10^{-0.5}\mathrm{Z}_{\odot}$ at redshift 5.7, 3. and 2.2 respectively.

The galaxies observed as LAEs exhibit a low metallicity independently of redshift or cosmic time. Additionally, LAEs also show the Z evolution present in the full galaxy population. Their metallicity distributions peak around $\sim 10^{-2.5}\mathrm{Z}_{\odot}$ at $z = 2.2$ and 3.0, while at $z = 5.7$ it peaks at $\sim 10^{-2.75}\mathrm{Z}_{\odot}$. There also differences between the Thin Shell and Galactic Wind models at low redshift. At $z = 2.2$ and 3.0 the Thin Shell predicts a higher number of LAEs with $\mathrm{Z} \sim 10^{-1.25}\mathrm{Z}_{\odot}$. Finally, very little differences are found between the LAE samples including RT in the IGM and excluding it.

4.5 LAE IGM properties.

In this section we study how large scale properties of the IGM affect our full RT LAE samples (FLAE). Later, in §4.7, we compare our results with previous works.

We analyze the Lyα IGM escape fraction computed as the ratio of the observed Lyα emission and the total Lyα flux escaping galaxies, i.e,

$$f_{\rm esc}^{\rm IGM} = \frac{{\rm L}_{{\rm Ly}\alpha}^{\rm Disk} \times f_{\rm esc}^{\rm Disk} + {\rm L}_{{\rm Ly}\alpha}^{\rm Bulge} \times f_{\rm esc}^{\rm Bulge}}{{\rm L}_{{\rm Ly}\alpha}^{\rm Disk} + {\rm L}_{{\rm Ly}\alpha}^{\rm Bulge}}, \tag{4.22}$$

as a function of the density ρ (Eq. 4.8), the IGM line of sight (LoS) velocity $\rm V_z$, the IGM gradient along the LoS of the velocity $\partial_z V_z$[5] and density $\partial_z \rho$. The gradients are computed from their respectively fields by a 3-point derivative method.

4.5.1 The Transmission of the IGM

In Fig.4.10 and Fig.4.11 we show the $f_{\rm esc}^{\rm IGM}$ behaviour of the Thin Shell and Galactic Wind models against the IGM properties at different redshifts. Overall, we find that the Thin Shell and Galactic Wind exhibit the same global trends. Since most of the Lyα photons are redshifted by the ISM, the IGM absorbs only a small fraction of the Lyα flux that escaped from galaxies. For both geometries, $f_{\rm esc}^{\rm IGM}$ is close to unity. The median $f_{\rm esc}^{\rm IGM}$ increases with the age of the Universe, as the IGM becomes more transparent to Lyα photons. We find that at $z = 2.2$ and 3.0 the IGM absorbs more photons in the Thin Shell model than in the Galactic Wind, while at $z = 5.7$ the absorption is comparable. In particular, for the Thin Shell, $f_{\rm esc}^{\rm IGM} \gtrsim 0.99$, ~ 0.96 , ~ 0.92 at redshift 2.2, 3.0 and 5.7 respectively. Meanwhile, the median $f_{\rm esc}^{\rm IGM}$ for the Galactic Wind rounds ~ 0.9999 , ~ 0.999 and ~ 0.92 at redshift 2.2, 3.0 and 5.7 respectively.

The differences between the Thin Shell and Galactic Wind are mainly due to their different Lyα line profiles. Our model is calibrated by fitting the luminosity function of our model with observations. There are two actors converting the intrinsic Lyα LF into the observed one: the RT in the ISM and in the IGM. As seen in Fig.4.8 the IGM effect in the LF is small at any redshift, which puts most (but not all) of the weight of the fitting in the ISM component. Roughly speaking, the shape of the observed LF determines the $\rm V_{exp}$ and $\rm N_H$ distributions (see Appendix) in our model through equations 4.1 and 4.2. Moreover, these distributions determine the properties of the Lyα line profiles and, therefore, the IGM absorption. Galaxies emitting more flux bluewards of Lyα would be more obscured than if they were emitting only redwards of Lyα. In practice, we find that the $\rm V_{exp}$ and $\rm N_H$ distributions make the Thin Shell to be more coupled to IGM at $z = 2.2$ and 3.0 while the opposite happens at $z = 5.7$,

---

[5] Throughout this paper we use the Einstein notation for partial derivatives, i.e., $\partial_z := \frac{\partial}{\partial z}$.

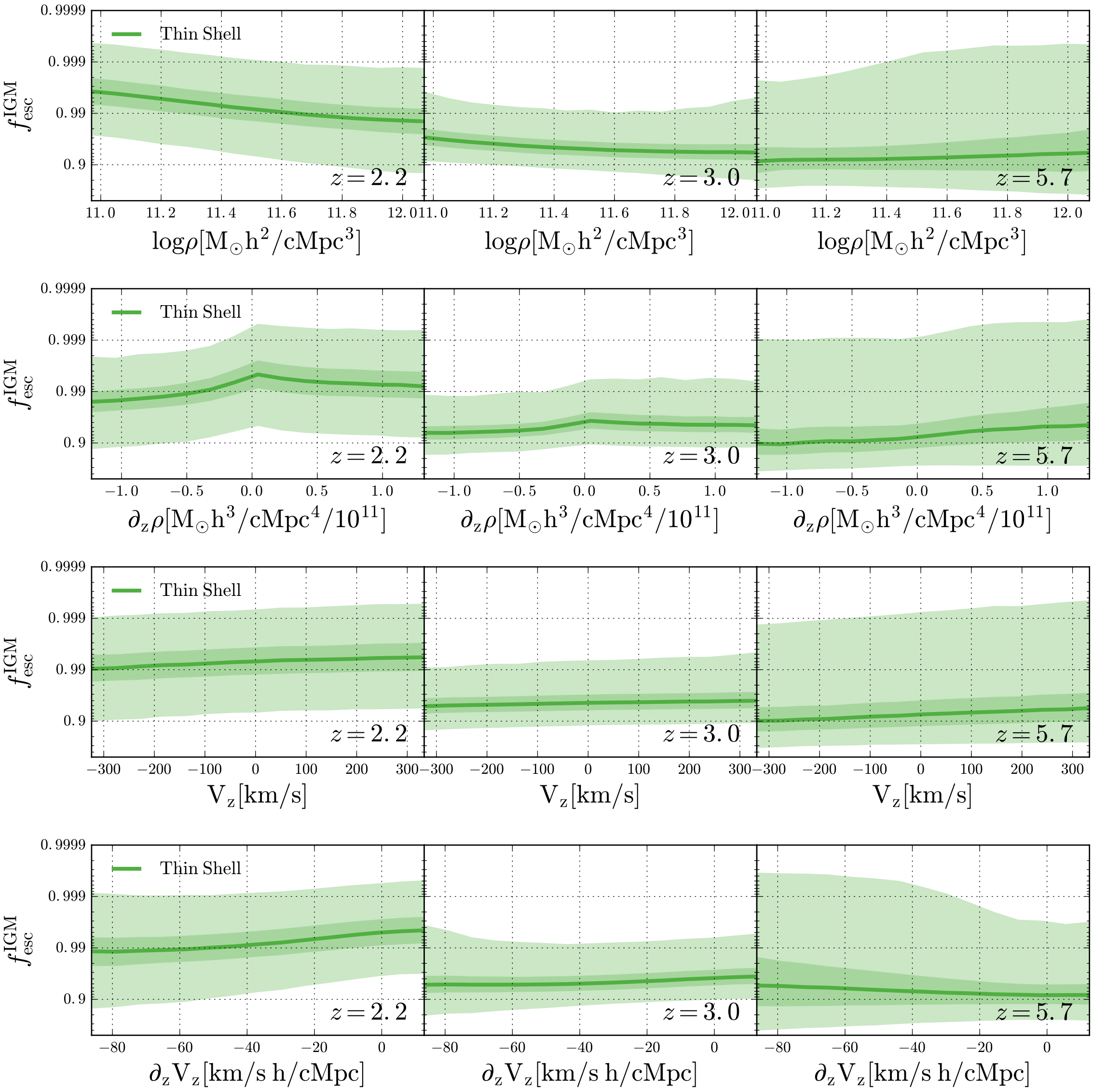


Fig. 4.10 IGM escape fraction for the Thin Shell as a function of the density, density gradient, velocity along the line of sight and velocity along the line of sight gradient (from top to bottom) for redshift 2.2, 3.0 and 5.7 (from left to right). The dark solid line shows the median, while the 1σ and 2σ values of the distributions are shown in dark and light shaded regions respectively.

where the Galactic Wind is more affected.

The median $f_{\rm esc}^{\rm IGM}$ varies over the dynamical range of the large scale IGM properties studied in this work. We note that the $f_{\rm esc}^{\rm IGM}$ variation is smaller than the dispersion around the median. However, the scatter is not caused by uncertainties, but by the

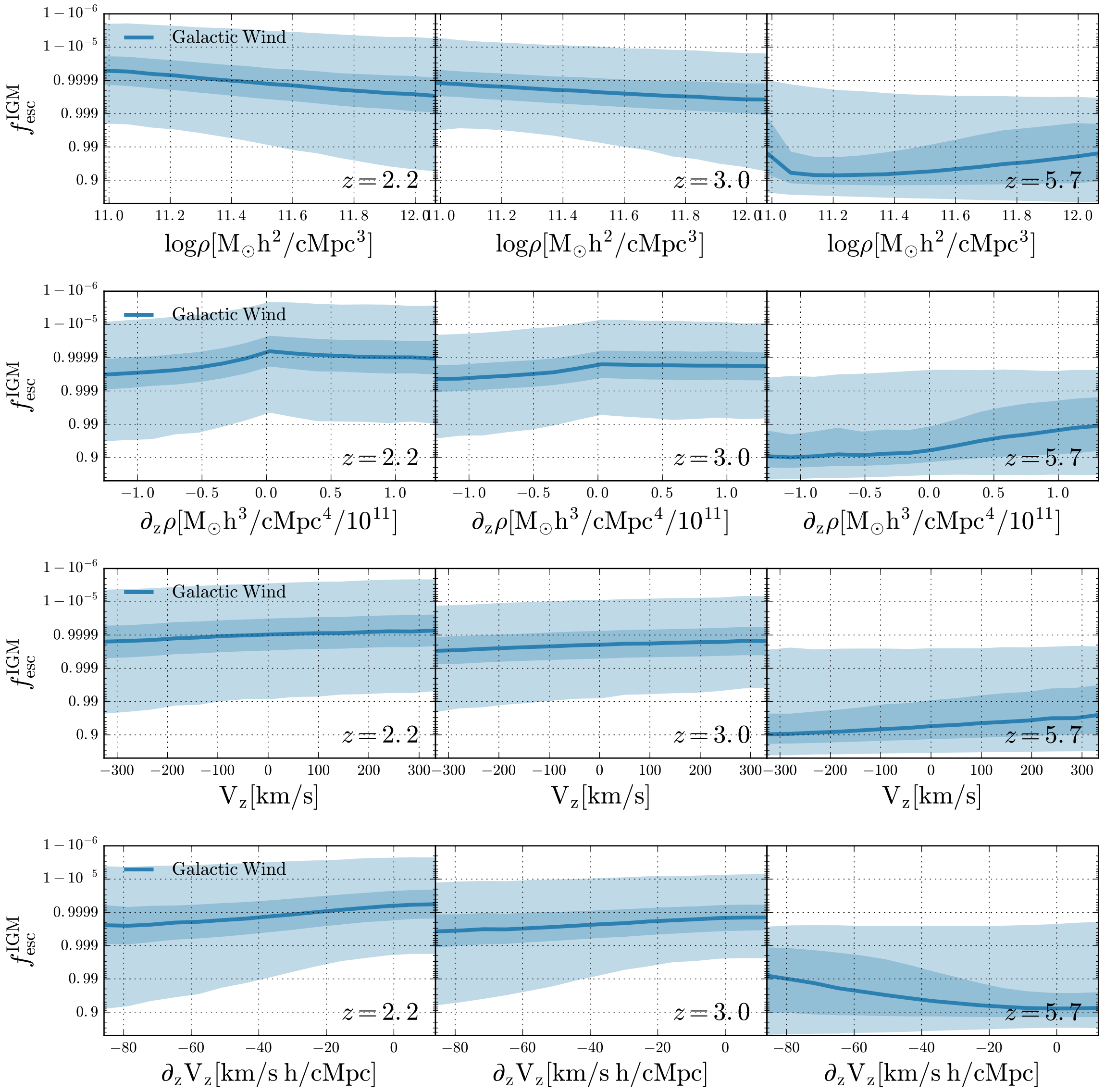


Fig. 4.11 Same as Fig. 4.10 but for the Galactic Wind geometry.

great diversity of combinations of ρ, $\partial_z\rho$, V_z and $\partial_z V_z$. This points in the same direction as Fig. 4.5, where the great variety of the IGM transmission curves is shown. These trends are statistically significant, i.e., they are not caused by noise or sample variance in our data set. In general, we find the same trends in the Thin Shell and Galactic Wind models. The strength of the correlations evolves with redshift. As the IGM becomes more transparent the trends become weaker. In fact, the lower the redshift, the weaker are the dependencies on $f_{\rm esc}^{\rm IGM}$. In particular, for the Thin Shell, the typical changes in $f_{\rm esc}^{\rm IGM}$ are $\lesssim 1\%$, $\sim 2\%$ and $\sim 5\%$ at $z = 2.2$, 3.0 and 5.7 respectively. Meanwhile, in the Galactic Wind modes, the $f_{\rm esc}^{\rm IGM}$ variations are $\lesssim 0.01\%$, $\lesssim 0.1\%$ and

$\sim 10\%$.

We find an anti-correlation between the IGM transmission and the local density. This is more apparent at $z = 2.2$ and dilutes at higher redshift. In particular, at $z = 5.7$ the median $f_{\rm esc}^{\rm IGM}$ is quite flat and the dispersion becomes greater at higher densities.

In general, $f_{\rm esc}^{\rm IGM}$ correlates with $\partial_z \rho$ at all redshifts. In addition to this trend, at redshift 2.2 and 3.0, $f_{\rm esc}^{\rm IGM}$ peaks at $\partial_z \rho = 0$, where the IGM transmission is slightly higher. However this peak is not present at $z = 5.7$.

We find a clear correlation between the IGM transmission and the IGM velocity along the line of sight at all redshifts. Galaxies in IGM regions moving towards the observer ($\rm V_z < 0$) suffer, statistically, greater absorption than galaxies moving away from the observer ($\rm V_z > 0$). The amplitude of the correlation augments towards higher redshifts, when the IGM becomes more optically thick. In particular, the IGM transmission variation in the $\rm V_z$ dynamical range rounds $< 1\%$, $\sim 1\%$ and $\sim 5\%$ in the Thin Shell at redshift 2.2, 3.0 and 5.7 respectively.

Finally, at $z = 2.2$, $f_{\rm esc}^{\rm IGM}$ correlates with $\partial_{\rm z} \rm V_z$. However, at $z = 3.0$ no correlation is found. At $z = 5.7$ we find a small anti-correlation that produces variations of $\sim 2\%$ in the IGM transmission in the $\partial_{\rm z} \rm V_z$ dynamical range.

4.5.2 The IGM-LAE coupling.

In Zheng et al. (2011) (from now on ZZ11) the authors present a model that features the LAEs coupling with the large scale properties of the IGM. The LAE models presented in this work exhibit similar trends to the relations found by ZZ11. In this section we discuss the origin of the IGM-LAE coupling, that matches ZZ11 interpretation for some IGM properties (ρ, $\partial_z \rho$, $\rm V_z$) while differs in others ($\partial_{\rm z} \rm V_z$).

IGM Density.

At low redshift, $f_{\rm esc}^{\rm IGM}$ anticorrelates with the density. This is caused by the column density along the line of sight between the galaxy and the distance where the IGM becomes transparent due to the Hubble flow. In regions of high density the amount of

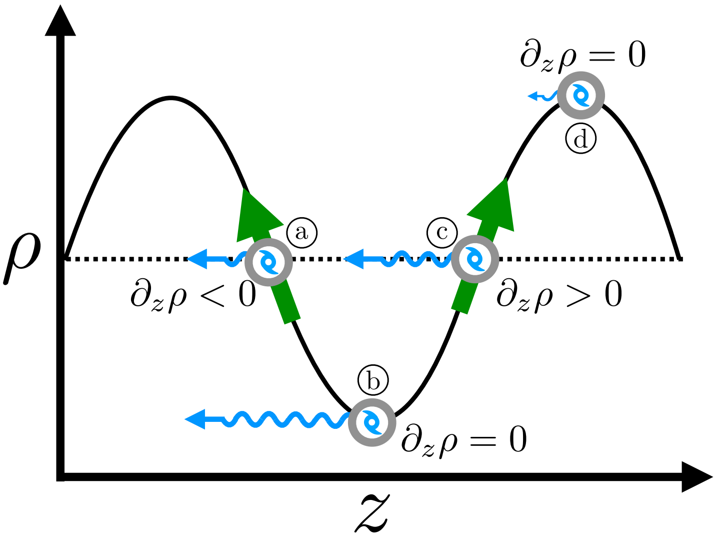


Fig. 4.12 In this figure the observer is always at the left edge. Galaxies are represented as grey circles with a spiral inside. The green arrows indicate the gradient. Lyα photons emitted toward the observer are shown with sinusoidal curves. The longer the photon arrows are, the higher the received flux and IGM transmission. Cartoon of how the ρ and $\partial_z \rho$ modify the Lyα transmission. The dark line represents a density fluctuation along the line of sight.

H$_{\rm I}$ that Lyα photons have to go through is greater than in environments with lower density. This causes a greater number of scattering events and absorptions, in the high density regions. Actually, this can be seen in Fig.4.6, where underdense regions exhibit a higher transmission than denser regions.

However, at $z = 5.7$ the trend reverses, and $f_{\rm esc}^{\rm IGM}$ is higher the greater the IGM density is. It is still true that the H$_{\rm I}$ column density is systematically higher in over dense regions. However, independently of the environment the IGM transmission at wavelengths bluer than Lyα is below 1%. This causes that the IGM selection effect on the density less important. At the same time, the impact of the IGM increases with redshift, causing strong selection effects on other IGM properties. Therefore, the correlation found between $f_{\rm esc}^{\rm IGM}$ and density is caused by the other IGM properties.

IGM density gradient.

Our models predict a correlation between $f_{\rm esc}^{\rm IGM}$ and $\partial_z \rho$. This trend is caused by the difference in H$_{\rm I}$ column density between the galaxy and the distance where the IGM becomes transparent. This scenario is illustrated in Fig.4.14. On one hand, for a source of Lyα photons at a fixed ρ, if the density gradient is positive (galaxy 'c'), then the IGM density decreases as the photons travel through it, allowing more photons to escape. On the other hand, in the case that $\partial_z \rho < 0$ (galaxy 'a'), the IGM density increases as photons go through the IGM, causing a higher number of lost photons.

Additionally, there are two main different regimes with $\partial_z \rho = 0$: i) the peak of overdensities (galaxy 'd' of Fig.4.12) with low IGM transmission due to the high column density; and ii) the bottom of underdensities (galaxy 'b') with high $f_{\rm esc}^{\rm IGM}$. The relative occupancy of these regions changes due to the RT in the IGM and ISM. LAE are preferentially observed in underdense environments. This is caused by two main reasons. First, the low IGM transmission in overdense regions. Second, the RT in the ISM prevents very massive galaxies (which would lie in overdense regions) to be observed as LAEs (see Fig.4.9). Therefore, most of the galaxies observed in environments with $\partial_z \rho = 0$ are in underdense regions. Meanwhile, the IGM transmission is higher the lower the density is. This causes the transmission peak found at $\partial_z \rho = 0$ for $z = 2.2$, 3.0.

IGM kinematics

When the relative velocity between the galaxy and the IGM is zero only photons bluewards of Lyα are absorbed, as seen in Fig. 4.5. However, if the IGM and the galaxy are moving towards each other, the Lyα line is blueshifted in the IGM rest frame, causing greater absorption. In the opposite scenario, where the galaxy and the IGM are moving away from each other, the Lyα photons are redshifted and escape the nearby IGM more easily.

The relative motion between the IGM and galaxies causes selection effects. On average, galaxies moving away from the observer are more likely to be observed as LAE. In Fig. 4.13 we list the different combinations of velocities between IGM and galaxies along the LoS.

- Case 1) Both the galaxy and the IGM are moving away form the observer. However, the galaxy velocity is greater. The IGM only absorbs bluer photon than Lyα.

- Case 2) Both the galaxy and the IGM are moving away form the observer. Additionally, the IGM moves faster than the galaxy. In this case, redder wavelengths than Lyα are absorbed.
- Case 3) The galaxy goes away while the IGM approaches the observer. Lyα photons are redshifted when they reach the IGM. Only bluer frequencies than Lyα suffer absorption.
- Case 4) Both the IGM and the galaxy approaches the observer. In this case the galaxy moves faster and the IGM absorbs redder wavelengths than Lyα.
- Case 5) Both the galaxy and IGM come towards the observer, but the IGM travels faster. The absorption only happens in bluer frequencies than Lyα.
- Case 6) The IGM moves away from the observer while the galaxy approaches it. The Lyα photons are blueshifted when they reach the IGM. Wavelengths redder than Lyα get absorbed.

Additionally, we have computed the relative abundance of these six scenarios in the galaxy population of `GALFORM`. We find that the cases 1) and 4) are equally probable, as well as 2) with 5) and 3) with 6). In particular, cases 1) and 4) constitute the 25% of the galaxy population each. Moreover, 2) and 4) represent 20% each, while 3) and 6) only 5% each.

On one hand, among the three scenarios where galaxies are moving away from the observer (cases 1, 2 and 3), in two of them (cases 1 and 3) the relative velocity between the IGM and galaxy is positive. In these cases the Lyα line profile is received redshifted in the IGM frame, causing low absorption. On the other hand, when the galaxy is approaching the observer (cases 4, 5 and 6) in two scenarios (4 and 6) the IGM sees Lyα blueshifted and absorbs Lyα photons. In addition, 60% of galaxies moving away from the observer are redshifted in the frame of the IGM. Meanwhile the 60% of galaxies approaching the observer are seen blueshifted by the IGM. This asymmetry causes that galaxies with $\mathrm{V_z} > 0$ (getting away from the observer) are more likely to be observed as LAEs.

IGM velocity gradient.

A further level of complexity is given by the cosmological velocity structure of the Universe (see Fig.4.2). The IGM is divided into regions of coherent motion of hundreds

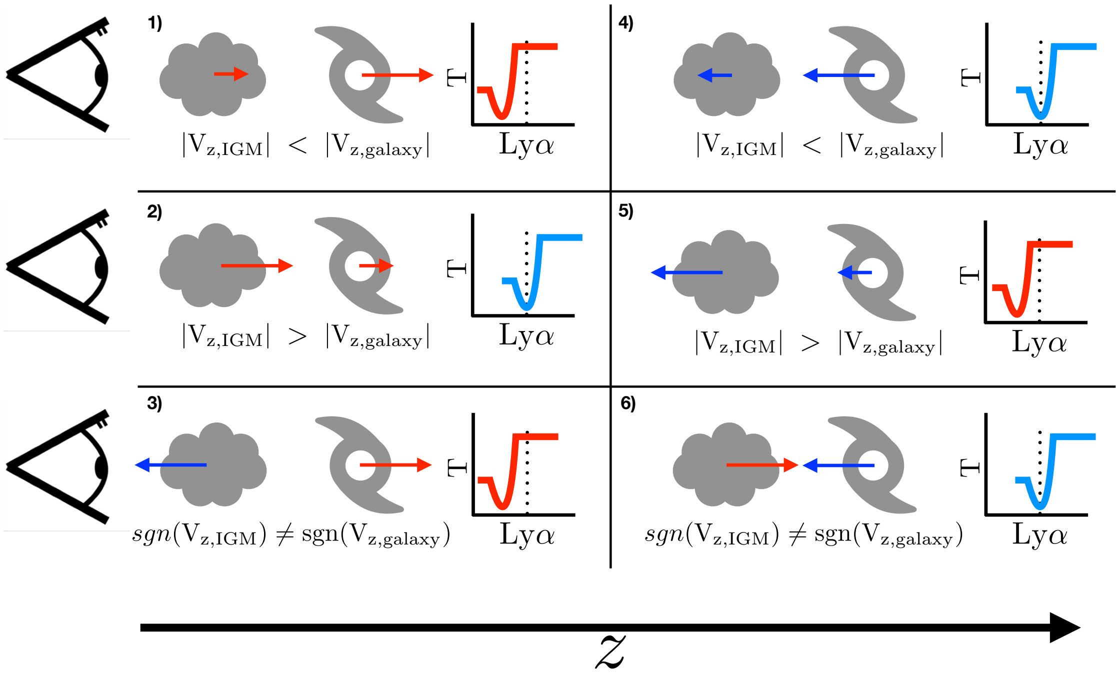


Fig. 4.13 Illustration of how the peculiar velocities of galaxies and the IGM change the Lyα observability. The observer is at the left edge. The figure is divided in six panels, one for each of the different combinations of IGM and galaxy motion (see §4.5.2). The clouds represent the IGM close to the galaxy along the LoS direction. The arrows are the velocity vectors of each component. The small plots are the IGM transmission in the galaxy rest frame versus wavelength. The dotted line within indicates the Lyα wavelength.

of cMpc that can be collapsing or moving away from each other. Additionally, the areas with $V_{\rm LoS} \sim 0$ between these regions are small (~ 5cMpch^{-1}), causing a great contrast of velocity on scales critical[6] to the IGM absorption.

There are two opposite effects controlling the IGM transmission dependence on ∂_zV$_z$. On the one hand, the transition between coherent motion regions facilitates the escape of Lyα photons emitted in clouds with $\partial_z V_z > 0$, as illustrated in the left panel of Fig.4.14. For example, if a given galaxy lies in the border of the IGM cloud with V$_{\rm z} > 0$ (bottom panel), Lyα photons are observed strongly redshifted in static (V$_{\rm z} \sim 0$) regions, and even more in clouds with V$_{\rm z} < 0$. In this case the Lyα escapes more easily. Meanwhile, in the opposite scenario, where the galaxy lies in a region with $\partial_z V_z < 0$, the neutral hydrogen increases the velocity towards the galaxy as the photon travels. This would result in a greater absorption. On the other hand, in

---

[6] Typically, the Hubble flow redshifts the Lyα line 1Å per $\sim$ 3cMpc at $z = 3.0$.

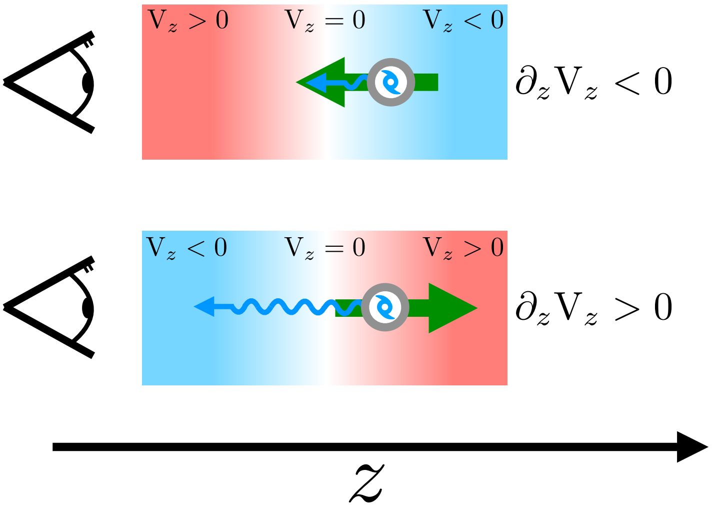


Fig. 4.14 In this figure the observer is always at the left edge. Galaxies are represented as grey circles with a spiral inside. The green arrows indicate the gradient. Lyα photons emitted toward the observer are shown with sinusoidal curves. The longer the photon arrows are, the higher the received flux and IGM transmission. Illustration of the $\partial_z V_z$ influence on the Lyα observability. The background color maps represent IGM regions with a velocity gradient. The velocity along the line is color coded from negative (blue) to positive (red).

regions with $\partial_z V_z < 0$ the Hubble flow effect is enhanced and photons escape more easily.

The Hubble parameter determines which of these effects prevail. If $H(z)$ is low, the typical distance that Lyα photons have to travel before the IGM becomes transparent is greater. In this case, this distance is compatible with the transition region between $V_z < 0$ and $V_z > 0$. Therefore, the Lyα transmission correlates with $\partial_z V_z < 0$. We find that in our model this happens at z=2.2 and 3.0. Meanwhile, if $H(z)$ is high, then the distance at which the IGM becomes transparent is smaller than the transition region. In this scenario regions with $\partial_z V_z > 0$ exhibit greater transmission. We find that this last scenario dominates $z = 5.7$ in our models.

4.6 The clustering of LAEs

In this section we analyze how the coupling of Lyα detectability and the IGM large scale properties modify the clustering of LAEs. To do so we compare the full RT LAE

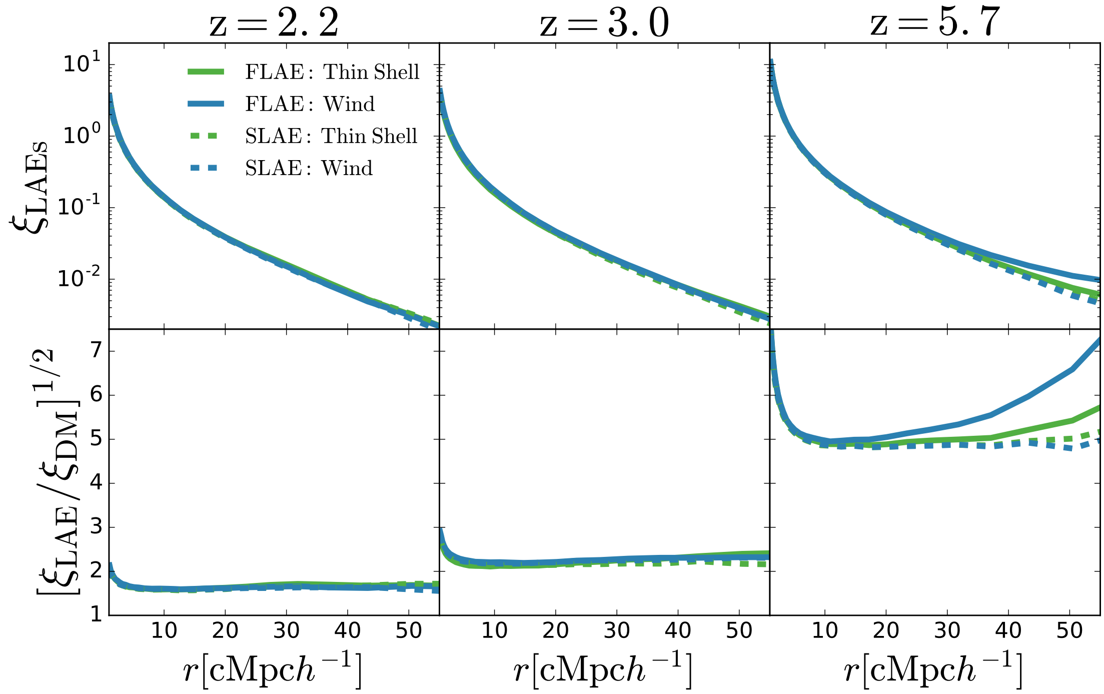


Fig. 4.15 **Top:** 2-point Correlation function in real space for different samples at different redshits (2.2, 3.0 and 5.7 from left to right). In dashed blue and green we show the SLAE Galactic Wind and Thin Shell respectively, while on the colors and solid line show the full RT models for the Galactic Wind and Thin Shell. **Bottom:** Bias of the different models as a function of distance.

sample (FLAE) and the SLAE. The SLAE populations inherit the mass function from the FLAE population. In contrast with the FLAE galaxies, by construction, the SLAE positions do not depend on the IGM transmission.

4.6.1 2-point 3D Correlation function.

In this section we analyze the clustering of our samples in real space. Note that similar result are found in redshift space.

In Fig. 4.15 we show the real-space 3D 2-point correlation function $\xi(r)$ and the bias computed as

$$b(r) = \left[\frac{\xi(r)}{\xi_m(r)}\right]^{1/2}, \tag{4.23}$$

where $\xi_m(r)$ is the matter correlation function.

By comparing LAE samples that include the full RT processes (FLAE) and their shuffled samples (SLAE) we can understand the IGM impact. We note that, by construction, our SLAE samples do not isolate the IGM effects but also includes assembly bias processes. However, we have checked that if we create shuffled samples from the PLAE populations (only RT in the ISM) the clustering measurements are identical, i.e., no evidence of assembly bias is found. Therefore, we attribute the differences in the clustering between FLAE and SLAE samples to only the IGM impact.

Overall, we find that the SLAE populations (both, Thin Shell and Galactic Wind geometry) behave in the same way in every redshift bin. Below $\sim$ 5cMpch^{-1} the bias of SLAE is not constant and decays with distance. From $\sim$ 5cMpch^{-1} on, their bias becomes scale-independent at a very similar value for both geometries. The bias increases with redshift. In detail, the bias of the SLAE on scales larger than 10cMpch^{-1} is $\sim$ 1.8, $\sim$ 2.4 and $\sim$ 5.2 at redshift 2.2, 3.0 and 5.7 respectively.

Meanwhile, the FLAE samples exhibit different behaviours at different epochs. On one hand, at redshifts 2.2 and 3.0, the FLAE and SLAE clustering exhibit the same trends and they are almost indistinguishable. Therefore, we find that the IGM does not shape the LAE clustering at these redshifts.

On the other hand, at larges distances, we find that the IGM increases the clustering of FLAEs at $z = 5.7$. At scales smaller than $\sim$ 5cMpch^{-1} the FLAE and SLAE clustering are identical. However, the FLAE samples including RT in the IGM exhibit a scale dependent clustering excess on scales $>$ 20cMpch^{-1}. The boost is present in both outflow geometries. However, its amplitude changes with the geometry. We find that the Galactic Wind geometry exhibit a more powerful boost than the Thin Shell. We attribute this to the greater coupling with the IGM in the Galactic Wind than in the Thin shell. In particular, the clustering at 50cMpch^{-1} is boosted a factor of $\sim$ 1.1 and $\sim$ 1.3 in the Thin Shell and Galactic Wind geometry respectively. We study the origin of the clustering boost in next section.

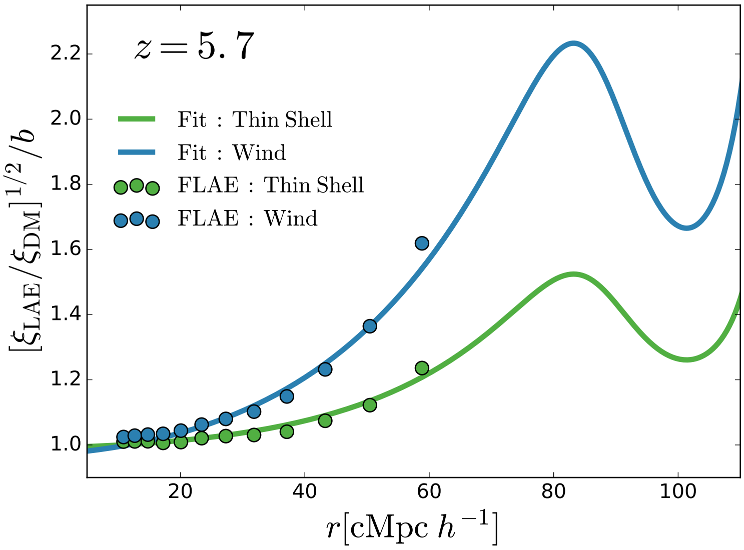


Fig. 4.16 Ratio between the bias as a function of distance and the median bias measured from 10 to 30cMpch^{-1} at redshift 5.7. In dots we show our simulation clustering while in solid curves the output of the MCMC analysis fitting the ZZ11 clustering description is illustrated. In blue and green we show the Galactic Wind and Thin Shell respectively.

4.6.2 The LAE clustering at large scales.

In this section we interpret our clustering measurements with the physical model presented in ZZ11. They described the overdensity field of galaxy population correlated with the large scale properties of the IGM as

$$\delta_g = b\delta_m \times \Big[1 + \tilde{\alpha}_1\delta_m + \tilde{\alpha}_2\frac{1}{aH}\partial_z \mathrm{V}_z + \\ + \tilde{\alpha}_3\frac{1}{aH}\left(\partial_x \mathrm{V}_x + \partial_y \mathrm{V}_y\right) + \\ + \tilde{\alpha}_4\frac{1}{aH}\frac{\mathrm{V}_z}{r_{\mathrm{H}}} + \tilde{\alpha}_5 r_{\mathrm{H}}\partial_z\delta_m\Big], \tag{4.24}$$

where a is the scale factor, δ_m is the overdensity of matter in the universe, $r_{\rm H}$ is a length scale set the coefficients dimensionless and the parameters $\tilde{\alpha}_i$ are free parameters that quantify the coupling with the IGM properties.

From this expression, the monopole of the galaxy power spectrum can be expressed in real space as a function of the matter power spectrum P_m :

$$P_0(k) = \left[\gamma_1^2 + \frac{2}{3}\gamma_1\gamma_2 + \frac{1}{5}\gamma_2^2 + \frac{1}{3}\gamma_3^2\right] b^2 P_m(k), \tag{4.25}$$

where

$$\gamma_1 = \left(1 + \frac{\tilde{\alpha}_1 - \tilde{\alpha}_3 f}{b}\right), \tag{4.26}$$

$$\gamma_2 = (\tilde{\alpha}_3 - \tilde{\alpha}_2)\,\beta, \tag{4.27}$$

and

$$\gamma_3 = \left(\tilde{\alpha}_4 \beta \frac{1}{k r_{\mathrm{H}}} + \tilde{\alpha}_5 \frac{k r_{\mathrm{H}}}{b}\right), \tag{4.28}$$

where f the growth factor and $\beta = f/b$. Meanwhile in redshift space, due to the redshift-space distortions (Kaiser, 1987) γ_2 is rewritten as

$$\gamma_2 = (1 + \tilde{\alpha}_3 - \tilde{\alpha}_2)\,\beta. \tag{4.29}$$

In this description the monopole (P_0) exhibits other additional dependencies on scale rather than P_m. There are two terms modifying the shape of the LAE power spectrum. First, a term proportional to $\tilde{\alpha}_4/k$ that enhances the clustering at large scales if the velocity field and LAE distribution are coupled. Second, a term proportional to $\tilde{\alpha}_5 k$ that amplifies the clustering in small scales if the LAEs are coupled to the gradient of density along the line of sight. Finally, the terms proportional to $\tilde{\alpha}_1$, $\tilde{\alpha}_2$ and $\tilde{\alpha}_3$ are scale independent and they only modify the clustering amplitude, leaving the monopole shape unaffected.

We perform a MCMC fit to Eq.4.25 using our model predictions at z=5.7 in real space. We assume that the bias b of the halo population hosting the FLAE samples is the same as their derived SLAE sample. In practice we compute b for each geometry outflow comparing the SLAE clustering to the matter clustering through Eq.4.23 and averaging between 10cMpch^{-1} and 30cMpch^{-1}. Additionally, we set $\tilde{\alpha}_1$, $\tilde{\alpha}_2$, $\tilde{\alpha}_3$ and $\tilde{\alpha}_5$ to zero, while $\tilde{\alpha}_4/\mathrm{r_H}$ is the only free parameter. In fact, we checked that the

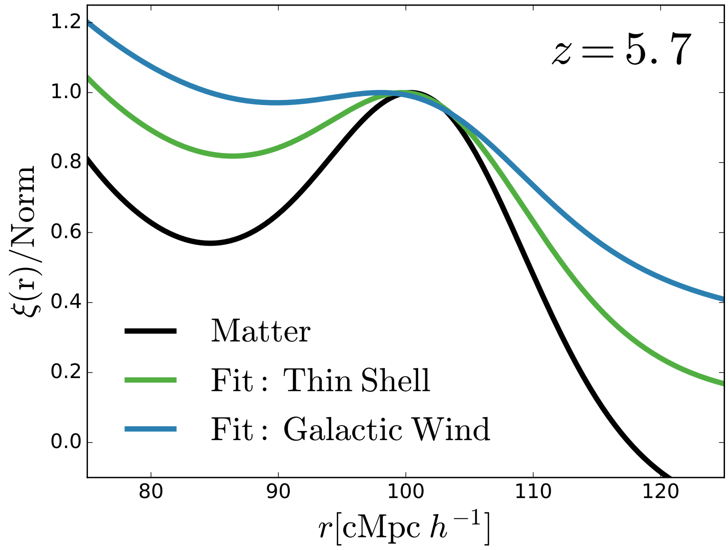


Fig. 4.17 Comparison between the 2PCF of our models and the matter 2PCF around the BAO peak. The clustering amplitudes are renormalized so that the maximun of the 2PCF between 90 and 110cMpch^{-1} match unity.

Table 4.3 Parameters of the the ZZ11 analytic expression (Eq.4.24) to the clustering measurements of our model at redshift 5.7.

Redshift	Geometry	b	log $\tilde{\alpha}_4/r_H$
$z = 5.7$	Thin Shell	4.82	-0.44
	Galactic Wind	4.87	-0.2

quality of the fit does not improve when the other parameters are included in the analysis. Finally, we restrict our analysis to the linear regime, i.e., between 5 and 60cMpch^{-1}.

The bias measurements and fitting results are listed in Table 4.3. In general, we find that the Thin Shell and Galactic Wind exhibit similar bias, as seen in Fig.4.15. Additionally, $\tilde{\alpha}_4$ is greater in the Galactic Wind than in the Thin Shell, indicating that the LAE model with the first one is more coupled to the IGM.

In Fig. 4.16 we compare our FLAE models (dots) with the MCMC output (solid lines). The ZZ11 description matches our simulations remarkably well for both outflow geometries. The extension of our models to larger scales highlights the complicated shape of the clustering at cosmological scales. In general, the ratio between the $\xi_{\rm LAE}(r)$ and $\xi_{\rm DM}(r)$ shows a hill at $\sim$ 80cMpch^{-1} and a valley at $\sim$ 100cMpch^{-1}, while it increases at even larger scales.

Also, in Fig. 4.17 we compare the matter 2-point correlation function (2PCF) with the analytic clustering description (calibrated with the MCMC analysis) of our Thin Shell and Galactic Wind FLAE samples. For the sake of a better comparison, we have renormalized the clustering of the different samples so that the 2PCF maximum between 90 and 110cMpch^{-1} matches in all cases. In the matter 2PCF we find the baryon acustic oscilation (BAO) peak, which is produced by the balance between gravity and pressure in the early Universe (e.g. Chaves-Montero et al., 2018). However, in the analytical description calibrated using our FLAE model, the shape of the BAO peak is distorted. In particular, it becomes broader in the Thin Shell geometry and even wider in the Galactic Wind geometry, as it is more affected by the IGM. Not only that, but also, the position of the maximum in our fits shifts by up to 1cMpch^{-1} and 3cMpch^{-1} respectively.

4.6.3 Clustering parallel and perpendicular to the line of sight.

In this section we study the clustering of LAEs in the perpendicular and parallel directions to the line of sight. In order to understand the role of the IGM we compare directly our FLAE and SLAE samples.

First, we make our analysis in real space to quantify the impact of the IGM on the LAE apparent spatial distribution. We perform our analysis at $z = 5.7$, where the coupling with the IGM is strong. Note that we find similar results at the other redshifts too. In Fig.4.18 we show the real space clustering of our LAE populations split in parallel and perpendicular to the line of sight of the IGM. We find that the LAE parallel and perpendicular clustering are symmetric. No angular dependence is apparently found in any of our models.

We also analyze the clustering of our LAE samples in redshift space. In order to convert from real space to redshift space we modify the galaxy positions with their peculiar motion along the LoS of the IGM (Z axis of our simulation) following

$$s = x_{\mathrm{Z}} + \frac{\mathrm{v}_{\mathrm{Z}}}{a(z)H(z)}, \tag{4.30}$$

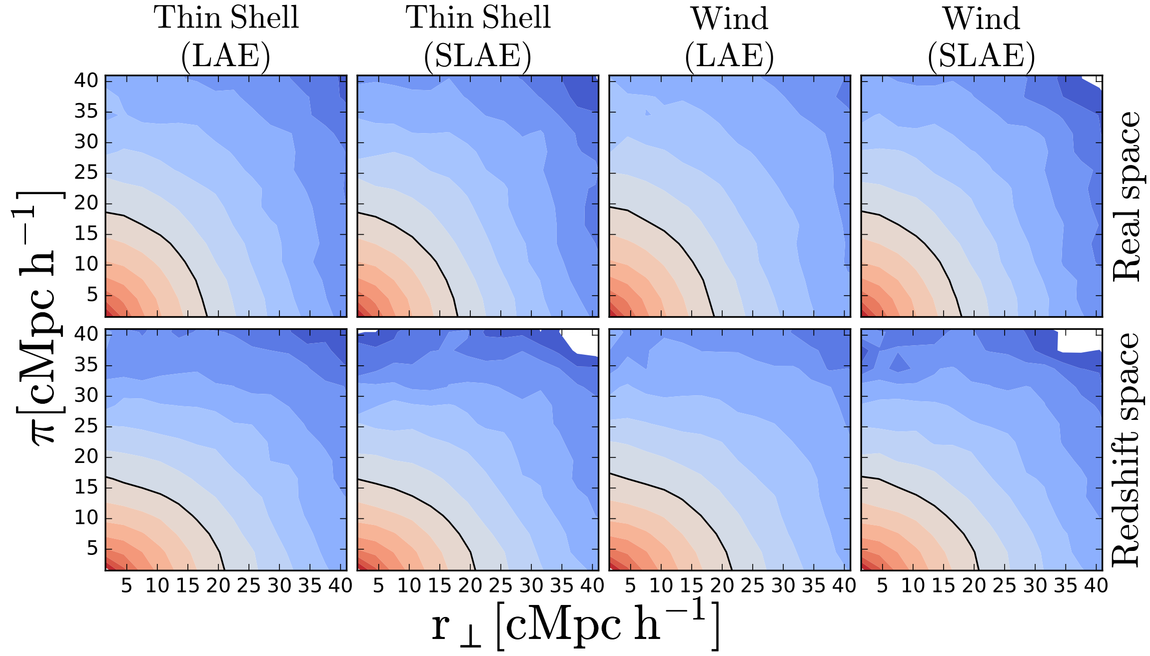


Fig. 4.18 LAE 2-point correlation function as a function of parallel (π) and perpendicular ($r_\perp$) distance to the line of sight in real space (top) and redshift space (bottom) at redshifts 5.7. From left to right we show the FLAE Thin Shell, the SLAE Thin Shell, the FLAE Galactic Wind and the SLAE Galactic Wind. The contours are curves of iso-clustering amplitude, divided in 15 bins from $\xi = 10^{0.5}$ to $\xi = 10^{-2.5}$ equispaced in logarithmic scale and are color-coded coherently. Additionally, the black curve indicate $\xi(\pi, r_\perp) = 0.01$.

where x_Z is the galaxy position along the LoS and V_Z is the galaxy peculiar velocity along the same direction.

In the bottom panel of Fig.4.18 we show the 2D clustering of our LAE samples at $z = 5.7$. In general, we find that in redshift space the clustering of all our models is suppressed along the LoS and conserved in the direction perpendicular to the LoS in every redshift bin. Additionally, the suppression along the LoS increases at lower redshift. In general, the peculiar velocity of galaxies points towards overdensities. Hence, the fluctuations along the line of sight are enhanced, which translates into the distortion of the clustering along the LoS.

Furthermore, we also analyze the clustering quadrupole of our FLAE and SLAE samples. In particular, the quadrupole is computed as

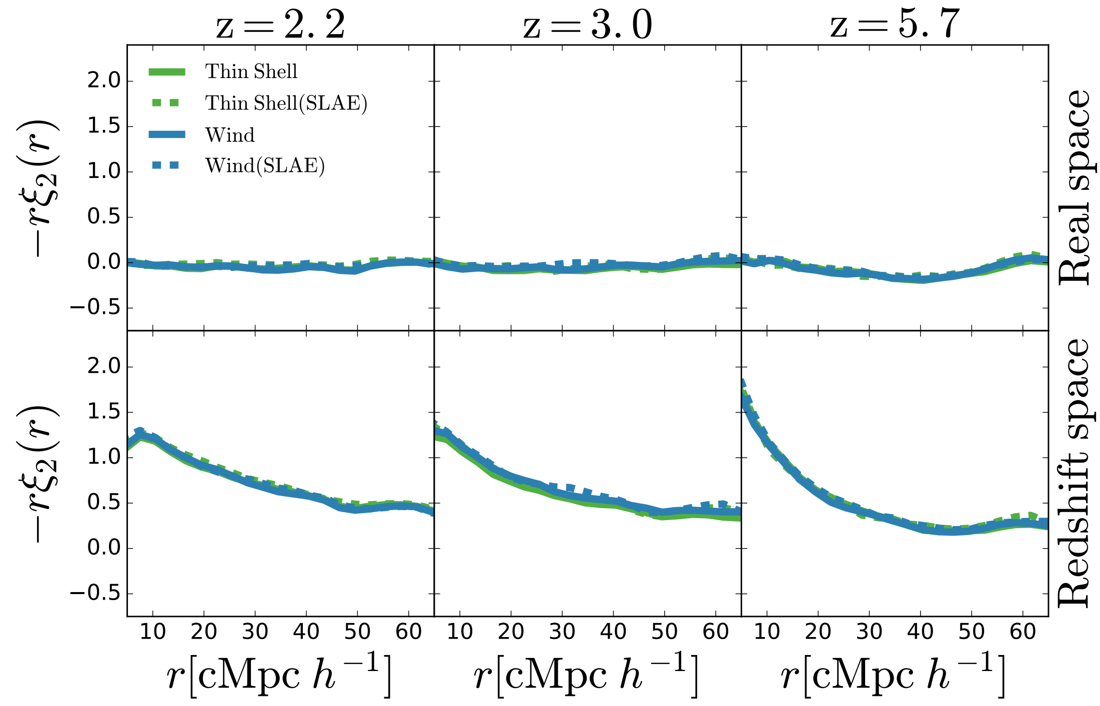


Fig. 4.19 Quadrupole of our different LAEs samples in real (top) and redshift space (bottom panels) at redshift 2.2, 3.0 and 5.7 from left to right. The Thin Shell and Galactic Wind are displayed in green and blue respectively. FLAE population are shown in continuum lines while dashed lines represent the SLAE samples.

$$\xi_\ell(r) = \frac{2\ell+1}{2}\int_{-1}^{1} d\mu\ \xi(r,\mu)\ \mathcal{L}_\ell(\mu) \tag{4.31}$$

where $\ell = 2$, $\mathcal{L}_\ell$ is the Legendre's polynomial of degree ℓ and $\mu = \cos\theta$, where θ is the angle between the line of sight and the sources.

In Fig.4.19 we show the quadrupole of the FLAE (continuum lines) and SLAE (dashed lines) samples at different redshifts in real space (top panels) and in redshift space (bottom panels). Overall, we find that at all redshifts the FLAE and SLAE quadrupoles are compatible in both, real and redshift space. Additionally, in real space (top panels) the quadrupole is null at all scales. This implies that the LAE-IGM coupling affects the parallel and perpendicular clustering in the same fashion, and no asymmetry is created. Then, in redshift space, we find a non-null quadrupole that evolves with redshift. The quadrupole becomes steeper on scales lower than 30cMpch^{-1} at higher redshifts.

4.7 Discussion.

In this section we compare our results to other theoretical works and current LAE observations.

4.7.1 Comparison with previous theoretical works.

In ZZ11the authors studied the coupling of the LAE observability and the IGM large scale properties. Later, Behrens et al. (2017) (from now on BC17) made the same analysis with similar techniques implemented in a simulation with higher spatial resolution. In this context, we compare our approach and results to these works.

Simulations

ZZ11 used a hybrid approach, where first a high resolution dark matter only N-body simulation was evolved while, in the fly, the hydrodynamic physics were run at lower resolution. The size of their simulation was $(100\mathrm{cMpc}h^{-1})^3$ with a constant spatial resolution of the neutral hydrogen density field close to $0.13\mathrm{cMpc}h^{-1}$. For further detail about the simulation we refer the reader to Zheng et al. (2010).

Meanwhile, BC17 implemented the IGM RT in the Illustris simulation (Nelson et al., 2015), a high resolution full hydrodynamic simulation using an adaptive mesh refinement approach. BC17 rebinned Illustris to a uniform grid of resolution $2\mathrm{cKpc}h^{-1}$. The size of this simulation was $(107\mathrm{cMpc}h^{-1})^3$. There are further information see their original work (Behrens et al., 2017).

Neither ZZ11 or BC17 were able to resolve the complex ISM structure due to its sub-kpc nature. Additionally, the volume of their simulation was not large enough to trace the large scale variation of the velocity field. In fact, their simulation could be enclosed within one of the many coherent motion regions populating our simulation (see Fig.4.2).

In contrast, our work uses a dark matter N-body simulation and implements the baryons in a post-processing flavour. Our resolution for the neutral hydrogen density is $0.2\mathrm{cMpc}h^{-1}$. However, our simulation size is $542.16\mathrm{cMpc}h^{-1}$, which translates into more than 125 times more volume than previous studies. This allow us to make accurate clustering predictions up to $\sim 60\mathrm{cMpc}h^{-1}$ and resolve the large scale structure of the velocity field of the Universe.

IGM radiative transfer methodology.

Both, ZZ11 and BC17 implement the RT in the IGM using post-processing Monte Carlo approaches similar to other works in the literature studying the Lyα RT in the ISM (e.g. Verhamme et al., 2006; Orsi et al., 2012; Gronke et al., 2016; Gurung-Lopez et al., 2018). In particular, in galaxy locations they generate photons in random directions and frequencies assuming a Gaussian line profile centered on Lyα. Then they track the photon's trajectory and changes in frequency. Finally, photons are collected and an $f_{\rm esc}^{\rm IGM}$ is computed for each galaxy comparing the number of emitted photons and the number of photons received within a given aperture centered on the galaxy position. Additionally, the Lyα luminosity emitted to the IGM by each galaxy is assumed to be directly proportional to the SFR (although in a different way in each work). For further details, we refer the reader to ZZ11 and BC17 original works.

Meanwhile, here we use a different approach also explored by Weinberger et al. (2018). We divide our simulation into thin cells along the line of sight and analytically compute the IGM transmission of each cells. Then we compute an absorption profile for each galaxy by summing all the small contribution of each cell. Finally, we convolve the IGM transmission with the Lyα line profile that depends on galaxy ISM properties such as the cold gas mass or the metallicity, as well as the Lyα luminosity emitted to the IGM (Gurung López et al., 2018).

IGM transmission

In comparison, ZZ11 and BC17 predict an IGM escape fraction well below that in our models. ZZ11 and BC17 RT approaches greatly overestimate the IGM absorption in comparison to our work for two main reasons:

1. The assumption of a Gaussian line profile centered on Lyα in contrast to much observational evidence supporting that the Lyα line profile is modified (and normally redshifted) by the ISM (e.g. Verhamme et al., 2008; Gronke, 2017; Sobral et al., 2018b). In this way, too much flux is put at bluer frequencies than Lyα, where the IGM is more efficient at absorbing photons (see Fig. 4.5).

2. The assumption that the $L_{Ly\alpha}$ emitted by a galaxy is directly proportional to the SFR. In fact, in our previous work (Gurung López et al., 2018) we found that the Lyα RT in the ISM breaks this relation. Their assumption places LAEs in more massive DM halos, hosted in denser environments, thus with lower median transmission (see Fig.4.6).

However, our model also has limitations. For example, we assume that every photon that interacts with the IGM is lost. In this way, we underestimate f_{esc}^{IGM} since we do not take into account photons that are scattered out of the LoS and thanks to consecutive scatters are sent back in the LoS again. The contribution of these photons to the received $L_{Ly\alpha}$ is predicted to be small (Zheng et al., 2010).

We note that at $z = 5.7$ the galaxy stellar mass distributions for our LAE samples truncates abruptly at $M_* = 10^7 M_\odot h^{-1}$. This is caused by the cut in `GALFORM` at this stellar mass imposed to ensure a good resolution of galaxies. This suggests that a fraction of the LAE population at $z = 5.7$ would inhabit galaxies with $M_* < 10^7 M_\odot h^{-1}$. However, as our galaxy population lacks these low mass galaxies, other more massive galaxies are selected as LAEs. Computing the precise number of LAEs that should host galaxies with $M_* < 10^7 M_\odot h^{-1}$ is challenging. As a simple calculation, we rescale the stellar mass distribution at $z = 3.0$ towards smaller M_* so that the peak of the distribution matches the one at $z = 5.7$. Then, the fraction of galaxies below the resolution limit is ~ 0.04. This hints that only a small part of the LAE population would lie in galaxies not resolved in our model. However, this limitation of our model might cause an overestimation of the clustering bias at this epoch (see below).

On one hand, ZZ11 found a strong correlation between the IGM and the LAE population. Their model predicted a relatively small dependence on V_z, ρ and $\partial_z \rho$ and a tight relation between $\partial_z V_z$ and f_{esc}^{IGM}, causing differences in f_{esc}^{IGM} greater than 1 order of magnitude across the $\partial_z V_z$ dynamical range.

In their model, the strong coupling between LAEs and $\partial_z V_z$ has dramatic consequences in the clustering of LAEs at $z = 5.7$. Their model predicts a scale independent enhancement in the clustering along the line of sight, which created an asymmetry between the clustering parallel and perpendicular to the LoS in real space. Moreover, in their model, the amplitude of this effect outpowers the Kaiser boost. Hence, even in redshift space the clustering along the LoS is more powerful than in the perpendicular direction.

On the other hand, BC17 claimed to find only a marginal coupling between the LAE and the IGM large scale properties at $z = 2.0$, 3.0, 4.0 and 5.85. Additionally, they studied the clustering of LAE samples and did not find any asymmetry between the directions parallel and perpendicular to the LoS clustering, nor any other strange clustering feature. Additionally, BC17 claimed ZZ11 results to be a consequence of poor spatial resolution. In particular, they recovered the strong selection on $\partial_z V_z$ after lowering their resolution to match ZZ11 simulation.

In this work we find a coupling between the Lyα observability and the large scale IGM properties. The amplitude of the coupling depends on redshift and on the outflow structure assumed in the ISM. Indeed, typical variations in $f_{\rm esc}^{\rm IGM}$ are $< 1\%$, $\sim 2\%$ and $\sim 5\%$ at $z = 2.2$, 3.0 and 5.7 respectively. In other words, we detect a LAE-IGM coupling but it is small, in particular at $z = 2.2$ and 3.0 where it has negligible effect on the clustering.

We do not find the dramatic asymmetry of clustering parallel and perpendicular to the line of sight (see Fig. 4.18) that ZZ11 measured in their original work. Our models do not predict any other clustering modification on scales where ZZ11 or BC17 simulations allowed them to measure the clustering accurately (~ 10cMpch^{-1}). Hence our clustering predictions are in agreement with BC17 and differ from ZZ11.

Our model predicts a strong boost of the LAE clustering at scales > 20cMpch^{-1}. This feature comes from the coupling with the IGM velocity field (see from eq.4.24 to eq.4.29). We attribute the detection of this effect in our work and the non-detection in BC17 to the difference in volume probed by the different simulations. Meanwhile, our simulation resolves the velocity large scale structure of the Universe, ZZ11 and BC17 simulations are small enough to be enclosed in one of the huge regions of coherent motion (see Fig. 4.2).

4.7.2 Comparison with current observations.

In this section we compare our model's clustering with current LAE clustering observations across cosmic time. In order to make a fair comparison we mimic the observations performed by Ouchi et al. (2010, 2018), Bielby et al. (2016) and Kusak-

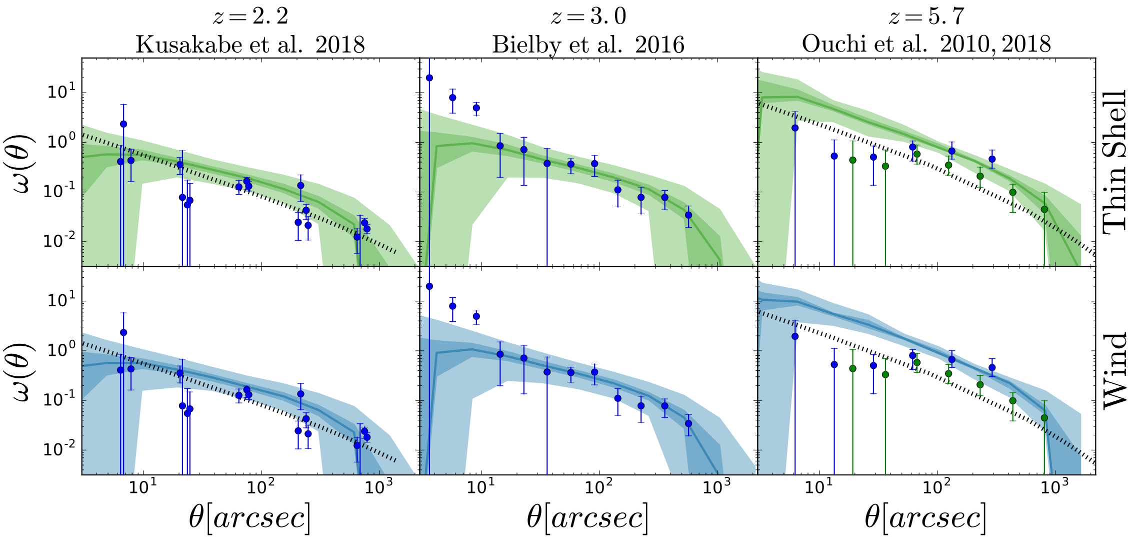


Fig. 4.20 Comparison between the clustering of our models (full RT Thin Shell, full RT Galactic Wind from top to bottom) and different observational data sets at redshift 2.2, 3.0 and 5.7 from left to right. At redshift 5.7, Ouchi et al. (2010) is shown in green dots while Ouchi et al. (2018) is plotted in blue dots. The solid colored line indicated the median angular 2PCF of our mocks while the 1σ and 2σ are shown as shade regions. The different observations are shown as dots. Finally, the dashed black line indicate illustrate the best fitting clustering model presented within the original works.

abe et al. (2018). These surveys select LAE candidates by combining narrow and broad band photometric filters. The main differences between the different surveys are the flux depth, the sky coverage and the redshift of LAE candidates. The different properties of the observations used in this work are listed in Table 4.5 and Table 4.4.

Here we summarize the mock construction and refer the readers to Gurung López et al. (2018) for a deeper explanation. In short:

1. We convert the galaxy coordinates to redshift space along the same line of sight (LoS) used to compute the IGM absorption (Eq.4.30).
2. We divide our box into rectangular subvolumes. The faces perpendicular to the LoS are squares with the same area as the survey sky coverage. The depth along the LoS is determined by the narrow band full width half maximum of each survey.

3. We make a rest frame equivalent width ($\mathrm{EW}_{0,\mathrm{cut}}$) and luminosity cut ($\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$). We choose the closest $\mathrm{EW}_{0,\mathrm{cut}}$ and $\mathrm{L}_{\mathrm{Ly}\alpha,\mathrm{cut}}$ to the survey cuts that best match the observed LAE number density. These cuts are listed in Table 4.5.

The mock properties are listed in Table 4.4. At $z = 2.2$ and 3.0 the survey volumes are small in comparison with our simulation and there are about 450 mocks in each redshift bin. Hence, we get a good estimate of the cosmic variance of these observations. However, at $z = 5.7$ the number of mock catalogs decreases to 18 and the constrain on cosmic variance is weaker. Additionally, the observed number density is well reproduced in our mocks. In particular, the median number of LAEs in each mock is within 1σ of the observed value.

In Fig. 4.20 we compare our full RT model mocks clustering with measurements at several redshifts. Overall, we find a good agreement with observations. In particular, at redshift 2.2, our RT models match very well observation at all scales. Meanwhile, at $z = 3.0$ the clustering is perfectly reproduced within the cosmic variance at angular separations higher than $\theta \gtrsim 10$ *arcsec*, while at smaller scales is underpredicted. Finally, at $z = 5.7$ we find that the full RT LAE samples exhibit a higher bias at all scales. This small disagreement is not caused by LAE-IGM coupling predicted by our model. If this was the case, only the large scales would disagree. In fact, the clustering excess here is partially caused by the overpredicted stellar masses at $z = 5.7$ in our FLAE samples, as described above (see Fig.4.9).

The characteristic excess of power in the clustering of FLAEs populations has been tentatively detected in Ouchi et al. (2018). For example, in Fig. 4.20 we show the LAE clustering measurements of SILVERRUSH (Ouchi et al., 2018) (in blue dots). The clustering at angular separations $> 10^2 arcsec$ is boosted. However, this trend is not found in Ouchi et al. (2010) (green dots). We attribute this to the different volume traced by surveys. On the one hand, Ouchi et al. (2010) only covered ~ 1 deg^2, an area small enough to fall inside one region of coherent motion (see Fig.4.2). On the other hand (Ouchi et al., 2018) explored ~ 13 deg^2 with contiguous patches of up to ~ 5 deg^2, big enough to begin to resolve these regions.

Table 4.4 Properties of our mock catalogs. In particular we list the redshift z, the redshift bin width δz, the size along and parallel to the line of sight ($L_{\|}$ and $L_{\perp}$ respectively), the Number of mock catalogs within our simulation volume and the median number of LAE in the catalogs for different models as well as the $\pm 1\sigma$ dispersion.

Authors	z	Δz	$L_{\|}$ (cMpc)	$L_{\perp}$ (cMpc)	N_{mocks}	$\langle N_{\mathrm{LAE}} \rangle$		
						Survey	Thin shell	Wind
Kusakabe et al. (2018)	2.2	0.0773	104.9	93.6	448	1248	1197^{+96}_{-86}	1193^{+96}_{-76}
Bielby et al. (2016)	3.0	0.0633	60.0	119.1	468	643	633^{+49}_{-52}	633^{+55}_{-47}
Ouchi et al. (2018)	5.7	0.0954	43.5	401.5	18	734	714^{+53}_{-28}	723^{+53}_{-51}

Table 4.5 Rest frame equivalent cut EW_0 and Lyα luminosity cut $L_{\mathrm{Ly}\alpha,\mathrm{cut}}$ in the different surveys and FLAE mocks.

Authors	$EW_{0,\mathrm{cut}}$[Å]			$L_{\mathrm{Ly}\alpha,\mathrm{cut}}$ [erg s^{-1}]		
	Survey	Thin Shell	Wind	Survey	Thin Shell	Wind
Kusakabe et al. (2018)	20.0	19.52	20.3	$1.62\ 10^{42}$	$1.47\ 10^{42}$	$1.77\ 10^{42}$
Bielby et al. (2016)	65.0	38.45	46.27	$1.62\ 10^{42}$	$1.32\ 10^{42}$	$1.52\ 10^{42}$
Ouchi et al. (2018)	20.0	20.06	20.06	$6.3\ 10^{42}$	$6.98\ 10^{42}$	$6.64\ 10^{42}$

4.8 Conclusions

We have created a cosmological model of Lyman-α emitter galaxies that includes Lyman α radiative transfer physics in both, the interstellar and intergalactic medium. The cosmological background is imprinted by the N-body simulation `P-Millennium`. Meanwhile, the semi-analytic model of galaxy of formation and evolution `GALFORM` populates the `P-Millennium` DM halos with galaxies. For the ISM transmission we used `FLaREON` (Gurung-Lopez et al., 2018), an open `Python` package based on a Monte Carlo RT code (Orsi et al., 2012) that predicts the Lyα line profiles and escape fractions of photons in outflows of different characteristics. Meanwhile, the RT in the IGM is implemented by computing the Lyα transmission at each position of our simulation. Our main conclusions are:

1. The RT in the ISM produces a strong selection effect over galaxy properties such as metallicity or SFR. Meanwhile, the RT in the IGM leave nearly unchanged the galaxy property distributions. In fact, LAEs tend to have a low-intermediate metallicity, moderate SFR and intermediate stellar mass. For further analysis on the galaxy properties of LAEs we refer the reader to Gurung López et al. (2018).

2. Our models predict that the Lyα IGM escape fraction depends on the large scale properties of the IGM such as the IGM density, motion, the density and velocity gradient along the line of sight, as first studied by Zheng et al. (2011). While at low redshift ($z = 2.2$ and 3.0) the correlations are weak and lead to variations in $f_{\rm esc}^{\rm IGM}$ of a $\sim 1\%$, it intensifies at higher redshifts ($z = 5.7$), reaching variation on $f_{\rm esc}^{\rm IGM}$ around $\sim 5\%$.

3. The level of coupling between the LAE distribution and the large scale IGM properties depends on the RT inside the ISM. Our model predicts that if the outflows driving the Lyα photons escape from galaxies have a Thin Shell geometry the coupling is greater than if it is driven by a Galactic Wind at low redshift. However, this the opposite is found at $z = 5.7$.

4. The IGM-LAE coupling can have an impact on their clustering. At redshift 5.7 the shape of the 2-point correlation function is modified, introducing extra power at scales larger than $\sim 20{\rm cMpc}h^{-1}$ in a scale-dependent fashion. Meanwhile, at $z = 2.2$ and z=3.0 we do not find any modification in the 2PCF, as the IGM-LAE coupling is too weak.

5. In order to study the LAE clustering at large scales we fit our model 2-point correlation function (2PCF) to an analytic clustering model including the IGM-LAE coupling (Eq.4.25) introduced by Zheng et al. (2011). The IGM-LAE coupling disrupts the LAE clustering and modify the shape of the 2PCF at the baryon acustic oscilation (BAO) peak scales, which becomes broader and modify the position of the maximum by $\sim 1\text{cMpc}h^{-1}$.

6. We have made a comparison between current LAE clustering measurements and our LAE models. Overall we find good agreement between our model and observations, even at $z = 5.7$. This suggests that a greater sky coverage is necessary to detect the clustering excess presented in this work. However, hints of a scale dependent bias at $z = 5.7$ can be found in the literature (Ouchi et al., 2018).

   In future works we plan to implement the model presented in this work to a larger simulation to determine the IGM-LAE coupling impact in the clustering at low redshifts. Additionally, we will also implement the physics of reionization to understand the clustering of LAEs during this interesting epoch.

4.9 Extra material

4.9.1 The $\mathrm{N_H}$, $\mathrm{V_{exp}}$ and τ_a distributions

In our model, the Lyα radiative transfer physics inside galaxies are integrated through `FLaREON`. `FLaREON` is open source code that predicts escape fraction and the emerging Lyα line profile from different outflow configurations among several gas geometries. In particular, `FLaREON` is based on pre-computed $\{\mathrm{N_H}, \mathrm{V_{exp}}, \tau_\mathrm{a}\ \}$ grids of the full radiative transfer Monte Carlo code `LyaRT` (Orsi et al., 2012). Then, different algorithms such as multidimensional interpolation are used to obtain the line profile and $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$. Hence, the high performance of `FLaREON` is limited to the space covered by the grid. In the following we study the fraction of LAEs in our model that fall within the `FLaREON` range.

In Fig. 4.21 we show the $\mathrm{N_H} - \mathrm{V_{exp}}$ distributions at different redshifts (solid colored lines) and the `FLaREON` accuracy boundaries (black lines). Over all, the Thin Shell and the Galactic Wind models behave likely and there is always some overlap between them. Additionally, as described above, there is little variation between the properties of the samples including the full RT (FLAE) and galaxies with only RT in in the ISM (PLAE). At low redshifts the $\mathrm{N_H} - \mathrm{V_{exp}}$ distributions are quite compact. In particular, at redshift

2.2 both geometries have a similar $\mathrm{N_H}$ distribution that peaks at $\sim 10^{20.5} cm^{-2}$, while the velocities of the Thin Shell ($\sim 10^{2} km\ s^{-1}$) are slightly above the Galactic Wind distribution ($\sim 10^{1.8} km\ s^{-1}$). Meanwhile, $z = 3.0$ the velocity rounds $\sim 10^{1.5} km\ s^{-1}$ and $\sim 10^{1.8} km\ s^{-1}$ in the Thin Shell and Galactic Wind respectively, while $\mathrm{N_H}$ are close to $\sim 10^{20.2} cm^{-2}$ and $\sim 10^{20.8} cm^{-2}$.

Finally, at $z = 5.7$ the distributions become broader. The velocities are higher and similar in both geometries, rounding $\sim 10^{2.4} km\ s^{-1}$, while the Thin Shell have higher $\mathrm{N_H}$ ($\sim 10^{19.8} cm^{-2}$) than the Galactic Wind ($\sim 10^{18.8} cm^{-2}$). We find that at that the fraction of LAEs outside the `FLaREON` range is less than a 3% in any of our samples.

In Fig. `FLaREON` we show the resulting τ_a distribution after calibration for our LAE samples. The τ_a distribution both geometries at $z = 2.2$ are narrow and centered at $\sim 10^{-1.5}$. At redshift 3.0 the distribution becomes slightly broader. Additionally, the Thin Shell exhibits lower dust optical depth ($\sim 10^{-1.8}$) than the Galactic Wind ($\sim 10^{-1.5}$). Meanwhile, at $z = 5.7$ the distribution become wider and the Galactic Wind have lower τ_a ($\sim 10^{-3.5}$) than the Thin Shell ($\sim 10^{-2.5}$). We also show the accuracy border of `FLaREON` at $\log \tau_a = 0.0$. We find that less than 1% of the LAE in our sample lie outside `FLaREON` τ_a dynamical range. All together we find that the fraction of LAEs outside `FLaREON` accuracy range is negligible.

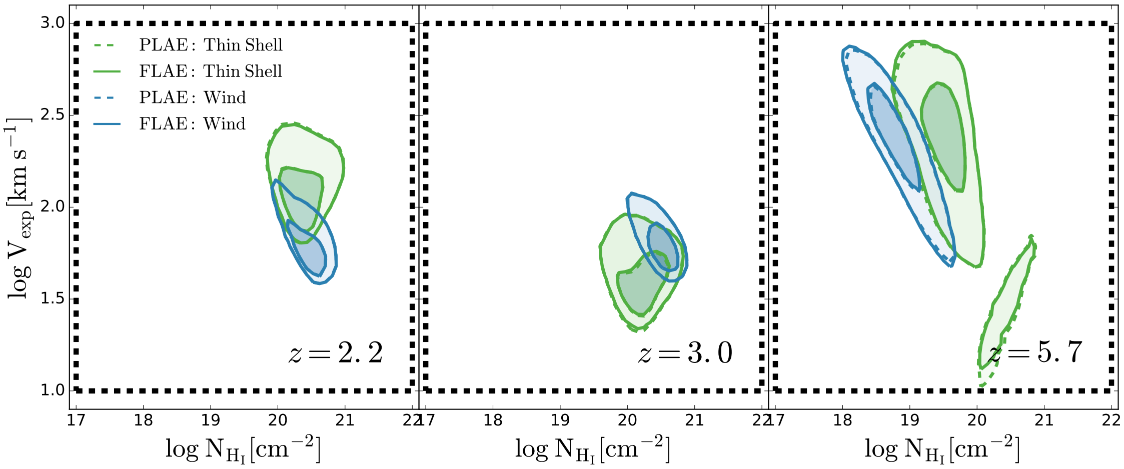


Fig. 4.21 Distribution of LAE in the V_{exp}-N_{H_I} space at redshift 2.2, 3.0 and 5.7 from left to right. The Thin Shell and Galactic Wind model are shown in green and blue. The dark countors represent the 1σ while the light the 2σ. The full RT models are shown in solid lines while the models with RT in the ISM but not in the IGM are plotted in dashed lines. In dashed black line we show `FLaREON` borders.

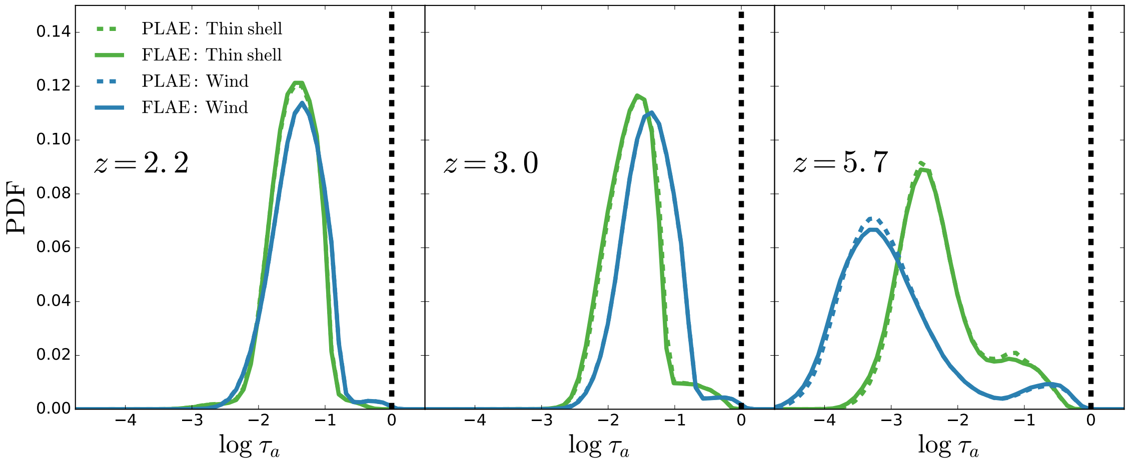


Fig. 4.22 Distribution of the dust optical depth our models including RT in the ISM at redshift 2.2, 3.0 and 5.7 from left to right. The color code of the lines is the same as in Fig.4.21.

Chapter 5

Imprints of reionization in Lyα selected samples

Abstract

The epoch of reionization is one of the most interesting time windows in the evolution of the Universe. Before the formation of the first stars and galaxies the intergalactic medium (IGM) was in a neutral state. Then, the energetic photons emitted by astrophysical sources progressively ionize the IGM. This transition is known as reionization. Here, we study the imprints that different histories of reionization leave in Lyα selected samples, by making use of the LAE model introduced in previous chapters. For this goal, we mimic the typical topology of the local fraction of neutral hydrogen during the epoch of reionization. This divides the simulation volume in two kind of regions: $\mathrm{H_I}$ clumps and $\mathrm{H_{II}}$ bubbles. Lyman-α emitters are mainly observed in $\mathrm{H_{II}}$ regions, while they avoid $\mathrm{H_I}$ clouds. We find that the history of reionization affects different properties of Lyman-α emitters. The amplitude of their luminosity function anticorrelates with the global fraction of neutral hydrogen. Moreover, we find that the clustering of this galaxy population depends on the stage and duration of reionization in different ways. First, the time of reionization modifies mainly the clustering amplitude, while the duration modifies the clustering shape. We conclude that Lyman-α emitters are a galaxy population that can help to determine the history of reionization with the help of future surveys.

5.1 Introduction

The Cosmic Dark Ages finished with the birth of the first stars. By this time the gas in the intergalactic medium (IGM) was cold and neutral. Then, hot OB stars and quasars started emitting photons capable of ionizing atomic hydrogen. This led to the progressive reionization of the Universe, affecting the topology of the local fraction of neutral hydrogen. At the beginning only small volumes around galaxies are ionized. However, as time passes these regions grow larger. Eventually, the H_{II} bubbles surrounding galaxies overlap with each other ionizing all the IGM. This scenario is known as inside-out reionization.

Additionally, different studies have found that the IGM topology during the reionization epoch may affect to the clustering of LAEs (e.g. McQuinn et al., 2007; Weinberger et al., 2019). This has motivated the use of this population of galaxies to put constrains in the history of reionization of the universe. Observational experiments such as SILVERRUSH (Ouchi et al., 2018) have already given an estimate of the global neutral fraction of the Universe using LAEs.

In this chapter we implement the LAE models presented in the previous chapter to the epoch of reionziation. In particular, we track the evolution of the global fraction of neutral hydrogen along the line of sight of our simulation assuming different histories of reionization. Then we study different properties of Lyα selected samples. This chapter is structured as follows: in §5.2 we explain in detail how we implement out model into the epoch of reionization. Then, in §5.3 we analyze the results of our model and we make our conclusions in §5.4.

5.2 Model description

In this chapter we implement the model introduced in Chapter 4 to the epoch of reionization. Our model is based on the N-body `P-Millennium` simulation, the semi-analytic model of galaxy formation and evolution `GALFORM`, `FLaREON` and the IGM absorption. Here we focus our analysis on the simulation snapshot produced at redshift 6.7.

5.2.1 Radiative transfer inside the interstellar medium

In order to implement the radiative transfer inside galaxies we use the same approach as described in § 4.2.1. In this chapter we only analyze the Thin Shell, as results are found to be similar for the Galactic Wind. In detail, we assign an outflow velocity $\mathrm{V}_{\mathrm{exp}}$, a neutral hydrogen column density N_{H} and a dust optical depth τ_a to each galaxy component (disk and bulge) of the galaxies produced by `GALFORM`, by using Eq. 4.1, 4.2 and 4.3 respectively. The free parameter of these expression are calibrated later in §5.2.4. In this way, using `FLaREON`, each component of each galaxy is assign a different escape fraction and a Lyα line profile that depends on their individual galactic properties.

5.2.2 Radiative transfer inside the intergalactic medium

The radiative transfer in the IGM is treated as in §4.2.2.

Fraction of neutral hydrogen

The main difference with the model presented in the previous chapter is the computation of the fraction of neutral hydrogen. During the epoch of reionization the global fraction of neutral hydrogen in the Universe $\langle \mathrm{X}_{\mathrm{H_I}} \rangle$ is close to unity. However, this does not mean that the local fraction of neutral hydrogen $\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x})$ is uniformly distributed close that value. Instead, the IGM is divided in two kind of regions: $\mathrm{H_I}$ clumps and $\mathrm{H_{II}}$ bubbles. The first regions are self-shielded against ionizing radiation, which causes that very energetic photons are not able to penetrate the gas. In this case the local fraction of neutral hydrogen is close to 1 all over the clump. In the second case, the $\mathrm{H_{II}}$ bubbles are regions that ionizing photons are able to penetrate. In this regions $\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x}) \ll 1$.

To implement these two different environments in our model we assume that regions below a certain phonotionization rate threshold Γ_{cut} are self-shielded. In these regions the local phonotionization rate $\Gamma(\overrightarrow{x})$ is attenuated an arbitrary factor. As a consequence, the flux of ionizing photons in self-shielded regions is negligible.

In short, in order to compute the local fraction of neutral hydrogen $\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x})$ in our model we

1. First: We compute $\Gamma(\overrightarrow{x})$ assuming that there are no self-shielded regions. For this we follow the methodology described in § 4.2.2. In detail, we compute the

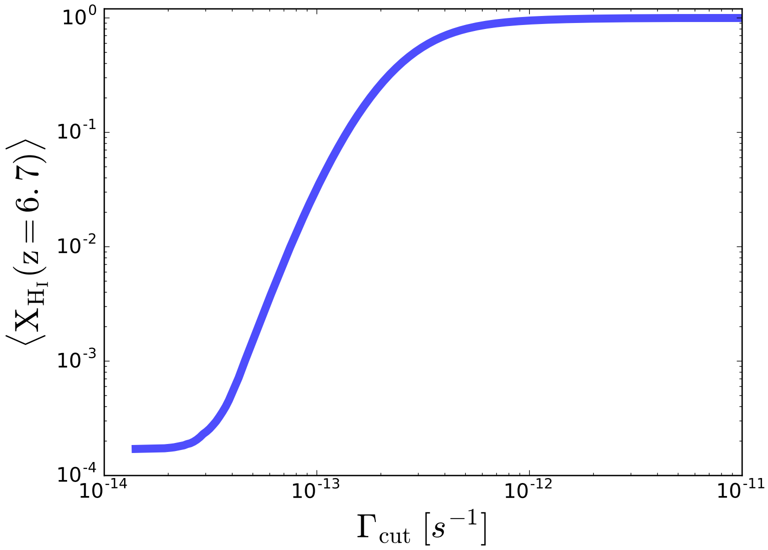


Fig. 5.1 Global fraction on neutral hydrogen in the whole simulation volume at $z = 6.7$ as a function of the threshold of ionizing rate.

contribution of each galaxy to $\Gamma(\overrightarrow{x})$ and then we fit the $\langle\Gamma\rangle$ predicted by Haardt and Madau (2012). This requires $\log K_a = 5.23$.

2. Second: We attenuate $\Gamma(\overrightarrow{x})$ in the regions where $\Gamma(\overrightarrow{x}) < \Gamma_{\rm cut}$.
3. Third: the local fraction of neutral hydrogen $X_{H_I}(\overrightarrow{x})$ is computed by using equation 4.13.

Then, the global fraction of neutral hydrogen in the Universe is computed as

$$\langle X_{H_I}\rangle = \frac{\sum X_{H_I}(\overrightarrow{x})\, n_H(\overrightarrow{x})}{\sum n_H(\overrightarrow{x})} \tag{5.1}$$

where the sums are over the simulation volume and $n_H(\overrightarrow{x})$ is the number density of hydrogen (neutral and ionized) computed as in the previous chapter (see §4.2.2).

In Fig. 5.1 we show the global fraction of neutral hydrogen in our simulation as a function of the threshold in the photoinization rate used. In general, the higher $\Gamma_{\rm cut}$ is, the higher is $\langle X_{H_I}\rangle$, as a grater volume of the simulation is considered shelf-shielded. In the high $\Gamma_{\rm cut}$ extreme $\langle X_{H_I}\rangle$ tends to unity. Meanwhile, in the low $\Gamma_{\rm cut}$ extreme $\langle X_{H_I}\rangle$ flats around 2×10^{-4} that corresponds with the global fraction of neutral hydrogen when the Universe is considered in the post-reionization epoch. Note, that if $\langle X_{H_I}\rangle$

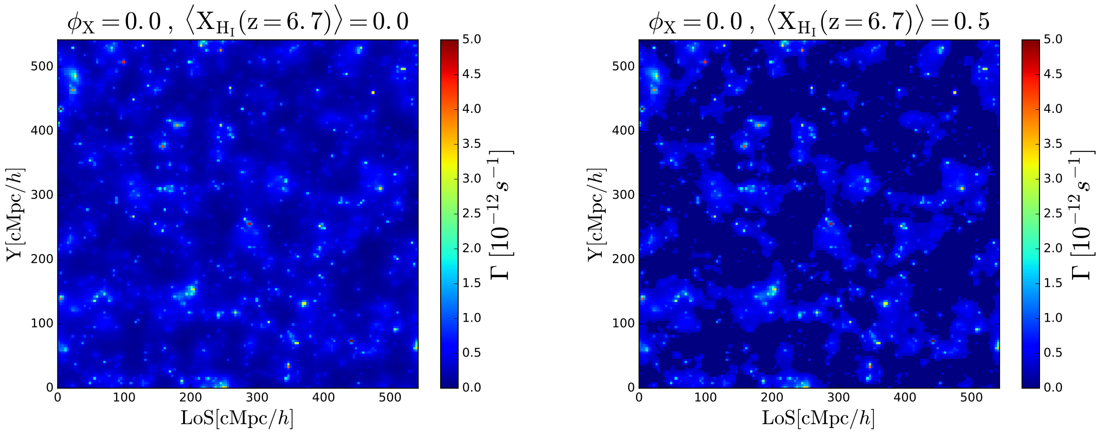


Fig. 5.2 Two examples of the photoionization field of our model in a slice of 3.6cMpc h^{-1}, where we have considered that all the simulation volume is at redshift 6.7. In the left we show a model without a photoionization threshold and in the right one where the global fraction of neutral hydrogen is 0.5.

was identically null, then Lyα photons and the IGM would not interact at all, making the transmission close to Lyα identically equal to unity.

Additionally, in Fig. 5.2 we compare two different computations of $\Gamma(\overrightarrow{x})$, one using a $\Gamma_{\rm cut} = 0$ (left), which will be denoted from now on as '$\langle X_{H_I} \rangle = 0.0$' (even if in reality $\langle X_{H_I} \rangle \sim 2 \times 10^{-4}$). In the right panel $\Gamma_{\rm cut}$ was chosen such as $\langle X_{H_I} \rangle = 0.5$. There are remarkable differences between both photoionization spatial distributions. For $\langle X_{H_I} \rangle = 0.0$ the distribution is smooth. Most of the volume has $\Gamma < 2 \times 10^{-12} s^{-1}$ and only a small regions exhibit strong ionizing flux. In contrast, when $\langle X_{H_I} \rangle = 0.5$, $\Gamma(\overrightarrow{x})$ is no longer smooth. The ionizing background is divided in two different environments: i) regions with high ionizing background which surrounds the regions with high Γ and ii) self-shielded regions which extend where $\Gamma(\overrightarrow{x}) < \Gamma_{\rm cut}$.

Although we make use of a very simple approach, the topology of $X_{H_I}(\overrightarrow{x})$ in our models resembles qualitatively to the one computed using more sophisticated techniques (Weinberger et al., 2018; Kulkarni et al., 2019).

Histories of reionization

In this chapter our main goal is to study how different histories of reionization affect Lyα selected samples. Up to now we have assumed that all the volume of our simulation was at the same redshift. However, from now on we will assume that the center of the simulation is exactly at redshift 6.7 and we let the redshift change along the line of sight. Given the size of our simulation the closest box face is placed at $z \sim 5.7$ and the farthest at $z \sim 7.9$.

We consider that the intrinsic $\Gamma(\overrightarrow{x})$ does not evolve along the simulation box. As can be seen in Fig. 4.3, $\langle\Gamma\rangle$ is quite flat from redshift 6 to 7 and it keeps flattening to higher redshifts. Additionally we also consider that the velocity field and the comoving dark matter density field are redshift independent. However, when we use the hydrogen density field in proper distance units we take into account the redshift dependence. In this way, the IGM is denser (in a proper frame) the further it is to the observer.

We use a very simple toy-model for the history of reionization. Our model assumes that $\langle \mathrm{X_{H_I}} \rangle = 1$ until the Universe starts to be ionized ($z > z_{\mathrm{X}=1}$) at a constant rate in redshift. After reionization is completed ($z < z_{\mathrm{X}=0}$), the global fraction of neutral hydrogen remains a few times 10^{-4} as already discussed. In particular, the redshift dependence of $\langle \mathrm{X_{H_I}} \rangle$ is

$$\langle \mathrm{X_{H_I}}(z) \rangle = \begin{cases} 1 & \text{when} \quad z > z_{\mathrm{X}=1} \\ z\,\phi_{\mathrm{X}} + \psi_{\mathrm{X}} & \text{when} \quad z_{\mathrm{X}=0} < z < z_{\mathrm{X}=1} \\ \sim 2\times10^{-4} & \text{when} \quad z < z_{\mathrm{X}=0} \end{cases}, \tag{5.2}$$

where ϕ_{X} is the rate at which the Universe gets ionized and ψ_{X} is chosen such as the global fraction of neutral hydrogen in the center of our box ($z = 6.7$) takes a certain desired value, i.e.,

$$\psi_{\mathrm{X}} = \langle \mathrm{X_{H_I}}(\mathrm{z} = 6.7) \rangle - 6.7\phi_{\mathrm{X}}, \tag{5.3}$$

where $\langle \mathrm{X_{H_I}}(\mathrm{z} = 6.7) \rangle$ is simply $\langle \mathrm{X_{H_I}} \rangle$ at the center of the box.

We build a total of 48 different reionization histories resulting from all the possible combinations of $\langle \mathrm{X_{H_I}}(\mathrm{z} = 6.7) \rangle = 0.0$, 0.1, 0.2 , 0.3, 0.4 , 0.5 , 0.6, 0.7 and $\phi_{\mathrm{X}} = 0.0$, 0.125, 0.25, 0.5, 1.0 and 2.0. The choice $\phi_{\mathrm{X}} = 0.0$ gives an scenario where the global fraction of neutral hydrogen does not evolve in our simulation. The value of $\phi_{\mathrm{X}} = 0.5$ roughly corresponds to the slope that is needed to start reionization at one edge of our simulation and finish it by the other edge, i.e., reionization take place between $z \sim 7.9$ and ~ 5.7. The other ϕ_{X} values were chosen as the quarter ($\phi_{\mathrm{X}} = 0.125$), the half ($\phi_{\rm X} = 0.25$), the double ($\phi_{\rm X} = 1.0$) and the quadruple ($\phi_{\rm X} = 2.0$) of this value.

In Fig. 5.3 we show the different histories of reionization used in this work. Our 48 reionization histories include very different cases. In some of them, a fraction of the volume of our simulation has $\langle {\rm X}_{\rm H_I}({\rm z})\rangle = 1$ and then reionization stars, lowering $\langle {\rm X}_{\rm H_I}\rangle$. In other cases the IGM at $z \sim 7.9$ is partially ionized as reionization is already ongoing. In other, reionization ends in the redshift range that our simulation covers and in other cases it finish later. Each panel of Fig. 5.3 shows all the different choices of $\phi_{\rm X}$ for a given $\langle {\rm X}_{\rm H_I}({\rm z} = 6.7)\rangle$, which causes that the different histories intersect at $z = 6.7$.

The photoionization rate topology during reionization

For each of the reionization histories we compute a different $\Gamma(\overrightarrow{x})$ field. In order to do so, we calculate $\langle {\rm X}_{\rm H_I}\rangle$ at each position along the line of sight of our simulation accordingly to a certain reionization history. Then, to each position we assign a different $\Gamma_{\rm cut}$ value that comes from the dependence on $\langle {\rm X}_{\rm H_I}\rangle$ shown in Fig. 5.1. Then, $\Gamma(\overrightarrow{x})$ is attenuated if it is lower than the assigned $\Gamma_{\rm cut}$ value.

In Fig. 5.4 we show $\Gamma(\overrightarrow{x})$ in the same simulation slice for all the possible combination of $\langle {\rm X}_{\rm H_I}(z = 6.7)\rangle$ and $\phi_{\rm X}$ studied in this work. There is a great variety of ${\rm X}_{\rm H_I}(\overrightarrow{x})$ topologies. In the cases where $\phi_{\rm X} = 0.0$, $\langle {\rm X}_{\rm H_I}\rangle$ is constant along the line of sight and the typical size of the $\rm H_{II}$ bubbles is maintained in the whole simulation. When $\phi_{\rm X}$ increases the evolution of $\langle {\rm X}_{\rm H_I}\rangle$ along the line of sight becomes apparent. In these cases the typical size of the $\rm H_{II}$ and $\rm H_I$ regions changes with line of sight distance. The $\rm H_{II}$ are smaller the further they are from the observer.

On one hand, in some cases where our simulation includes volume where reionization has not started yet. This corresponds to the cases where $z_{\rm X=1} \in (5.7, 7.9)$. On the other hand, there are cases in which $\langle {\rm X}_{\rm H_I}\rangle$ drops to ~ 0 up to a certain distance. Theses cases correspond with $z_{\rm X=0} \in (5.7, 7.9)$. Additionally, these two last scenarios are found together when $\phi_{\rm X}$ is big enough (e.g. $\phi_{\rm X} = 1.362$ and $\phi_{\rm X} = 2.724$) and the simulation is divided in two clear regions: before and after reionization. The transition between these regimes depends on $\phi_{\rm X}$. The higher $\phi_{\rm X}$ is, the faster it is. In the most extreme cases where $\phi_{\rm X} = 2.724$ reionization takes place in only $\sim 200{\rm cMpc}\ h^{-1}$.

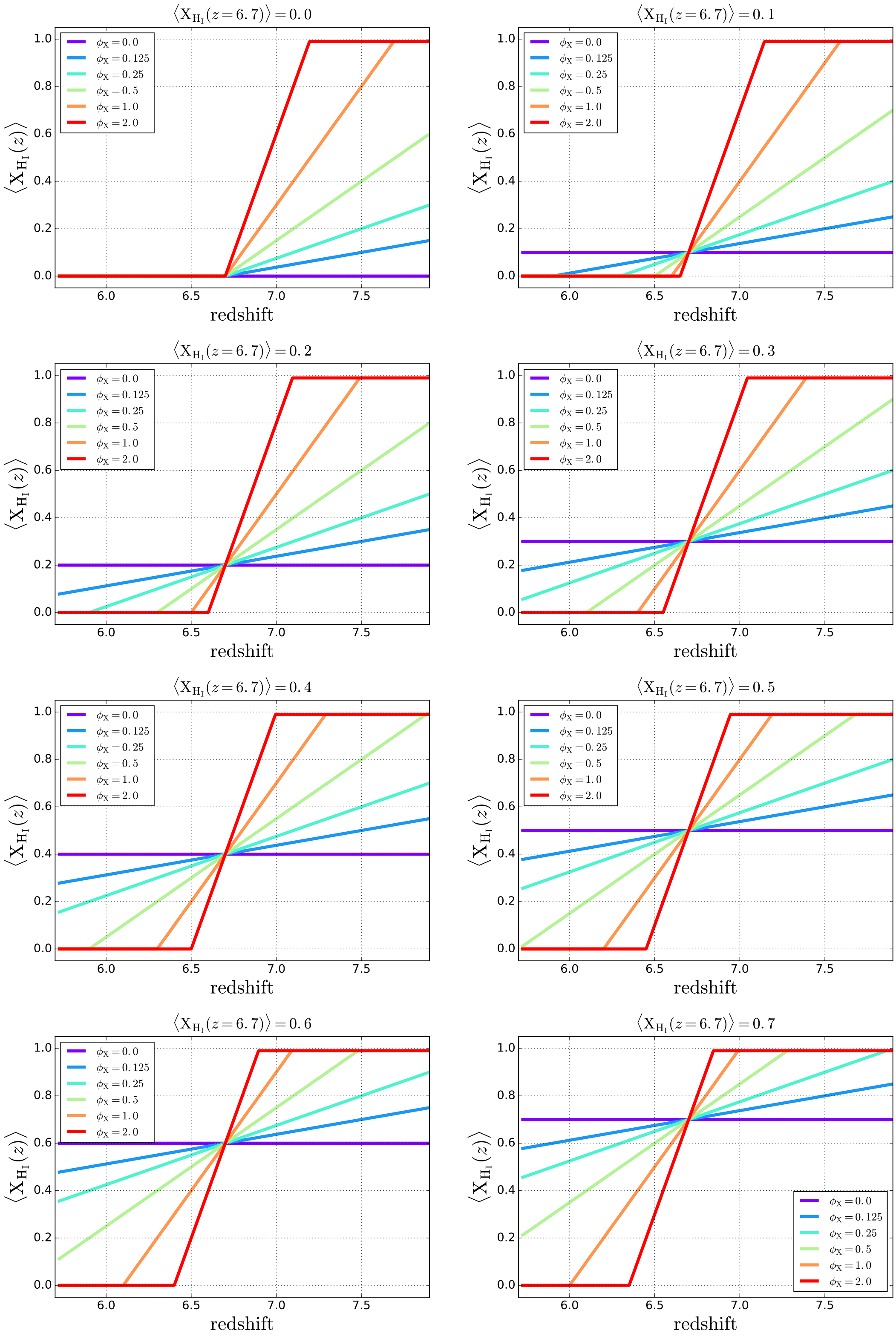


Fig. 5.3 All the reionization histories considered in this chapter. Each panel shows a fixed value of global fraction of neutral hydrogen at z=6.7, while the different colored lines shows a different values of ϕ_X as labeled.

Finally, for each reionization history, the local fraction of neutral hydrogen $X_{H_I}(\overrightarrow{x})$ is computed as before (using equation 4.13). In this way, the regions that were attenuated result in H_I clumps. Meanwhile, the other regions remain as H_{II} regions with a fraction of neutral hydrogen close to 2×10^{-4}.

5.2.3 IGM Transmission

In the model presented in this chapter we compute the IGM transmission as in the previous chapter (see § 4.2.2).

In Fig. 5.5 we compare the transmission curves of galaxies placed in H_{II} bubbles and H_I clumps. The original sample is composed by the 150000 galaxies with the highest sSFR, as representatives of emission line galaxies. Then this sample is split in two depending on whether the galaxy lies in a neutral or an ionized region. To make this comparison we used the history of reionization with $\langle X_{H_I}(z = 6.7)\rangle = 0.3$ and $\phi_X = 0.0$. The transmission inside the H_{II} bubbles is 1 at wavelengths bluer than Lyα, then it drops very fast close to Lyα and at redder wavelengths it is close to 0. Indeed, they look like the transmissions in the post-reionization epoch (see Fig. 4.5), where all the IGM is ionized. In contrast, the IGM transmission inside H_I clumps behave different. While the transmission at wavelengths redder than Lyα is also close to 0, the transmission at bluer wavelengths is lower than 1. In detail, the IGM absorption is so strong that the red wing of the Lyα-H_I cross-section causes that photons redder that Lyα are also scatter out of the line of sight. We have checked that the transmission keeps increasing towards redder wavelengths until the cross-section becomes negligible and $T = 1$.

In this way, galaxies lying in H_I clumps will have a huge attenuation of the Lyα line. Most of the galaxies that are intrinsically LAE are not observed as such. Meanwhile, if the Lyα luminosity is great enough the galaxy might still be observed as LAE. This depends strongly on the Lyα line profile.

5.2.4 Model calibration

Our implementation of the RT inside galaxies has some free parameters that control the Lyα escape fraction from galaxies and the Lyα line profile injected into the IGM through V_{exp} and N_H. These properties are computed by using the recipes in Eq. 4.1

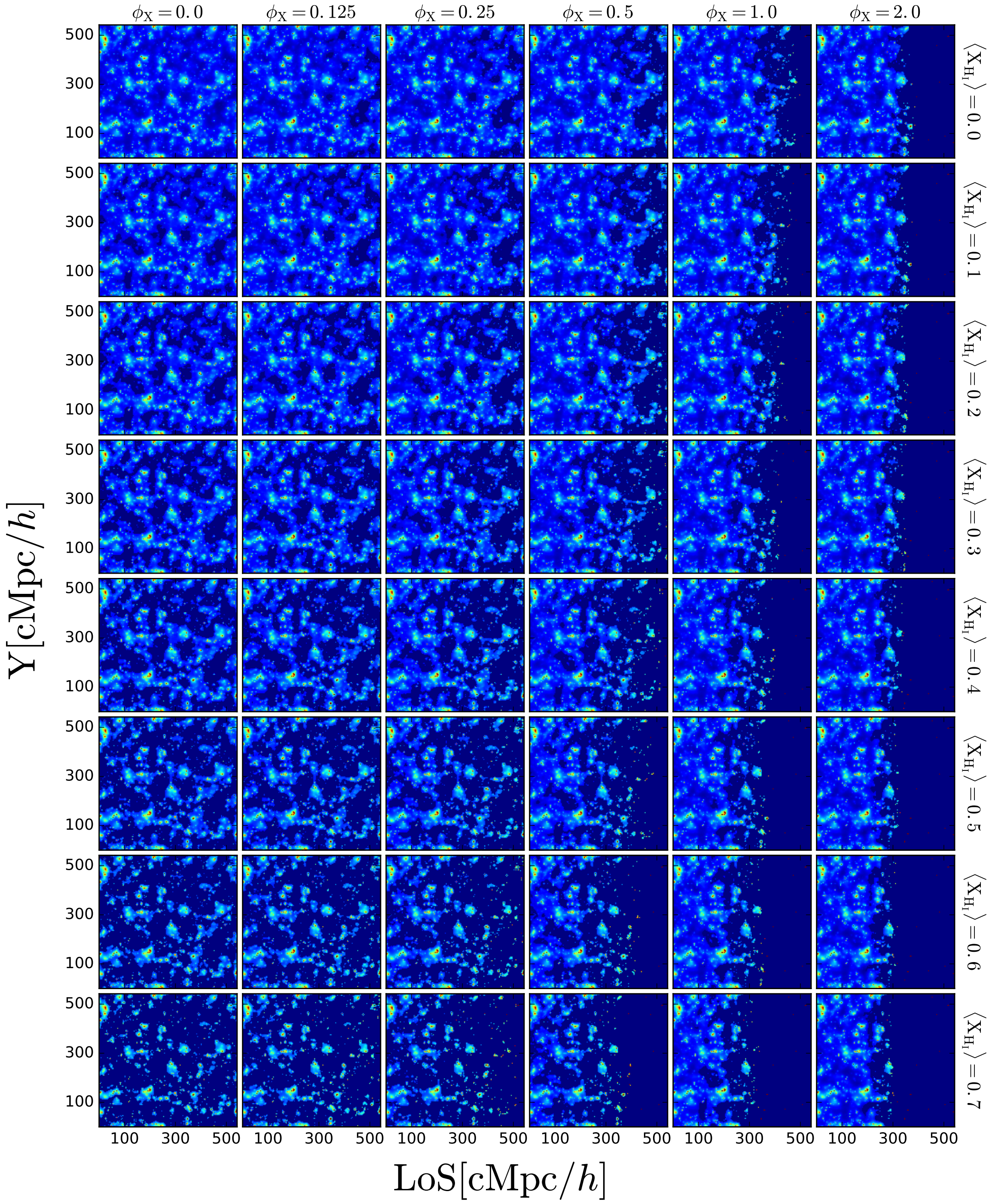


Fig. 5.4 Photonization rate field for each of the histories of reionization considered in a slice of 3.6cMpc h^{-1}. The color map is the same as in Fig. 5.2. Columns have a fixed value of ϕ_X and rows a fixed value of global fraction of neutral hydrogen at redshift 6.7. The horizontal direction represents the line of sight.

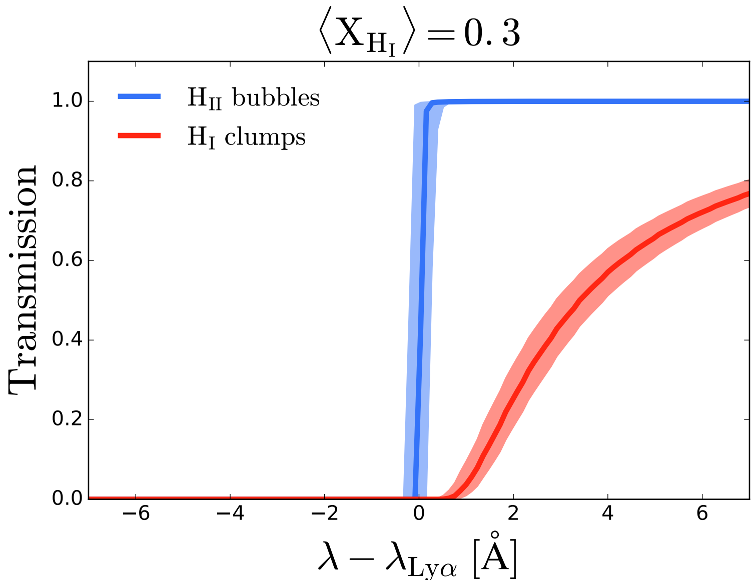


Fig. 5.5 Comparison of the IGM Lyα transmission during the epoch of reionization. In blue we show regions inside H_{II} bubbles and in red, regions inside H_I clumps. This data were taken from the $\langle X_{H_I}(z = 6.7) \rangle = 0.3$ and $\phi_X = 0$ model.

and 4.2. The calibration of these free parameters is made as in the previous chapter (see§ 4.3.1). In short we set the value of the free parameters by fitting the observed luminosity function of LAEs at $z = 6.7$ Ouchi et al. (2008); Konno et al. (2018). For each combination of free parameters we compute for each galaxy the Lyα luminosity after the ISM and IGM absorption by using Eq. 4.18.

To calibrate our model we make use of a $X_{H_I}(\overrightarrow{x})$ that assumes that i) all the simulation is at $z = 6.7$, i.e, there is no line of sight redshift dependence and ii) the global fraction of neutral hydrogen in the Universe is $\langle X_{H_I}(z = 6.7) \rangle$. Our choice of the $X_{H_I}(\overrightarrow{x})$ used in the calibration is motivated by three factors:

1. First, the specifics of the LAE surveys performed by Ouchi et al. (2008) and Konno et al. (2018). These experiments used a photometric narrow band to detect Lyα emission. The FWHM (full width half maximum) of their narrow band allowed them to detect LAEs at $z \sim 6.7$ in a redshift window of $\Delta z = 0.1078$, which translates in roughly $L_{\|} = 41$cMpc along the line of sight. There is the possibility that the volume of the Universe they measured had different $\langle X_{H_I}(z) \rangle$. Considering that $\langle X_{HI} \rangle$ is constant in a range of $\Delta z = 0.1078$ is appropriate for our models with low ϕ_X, where the typical Δz necessary to complete reionization is greater than 1.5. However, in our models with a fast history of reionization this

Table 5.1 Free parameters as defined in equations 4.1 and 4.2 after the calibration with the luminosity function observed by Ouchi et al. (2008); Konno et al. (2018).

redshift	Geometry	$\langle X_{HI} \rangle$	$\log \kappa_{V,disk}$	$\log \kappa_{V,bulge}$	$\log \kappa_{N,disk}$	$\log \kappa_{N,bulge}$
$z = 6.7$	Thin Shell	0.3	4.77	4.05	-15.47	-12.11

is not longer valid since the typical time to completely reionized in the Universe is $\Delta z \sim 0.2$. However, as we will see later (Fig. 5.10), for fixed value of $\langle X_{H_I} \rangle$, the LAE luminosity function does not depend strongly on ϕ_X (specially if $\langle X_{H_I} \rangle$ is high). We do not expect that this variation affects heavily to our calibration.

2. Second, by assuming that all the box is at 6.7 we increase by a factor of ~ 20 the volume available to compute the LF and calibrate our model. This reduces the cosmic variance that volumes replicating the depth of the surveys(i.e., $\Delta z = 0.1078$) would suffer.

3. Third, the Planck mission (Planck Collaboration et al., 2016) determined that $\langle X_{H_I}(z = 6.7) \rangle \sim 0.3$. As we will see later, the LF amplitude depends strongly on $\langle X_{H_I} \rangle$. Variations on the $\langle X_{H_I} \rangle$ value used in the calibration would result in different values for the free parameters.

The values of the calibrated free parameters are listed in Table 5.1. Additionally, in Fig. 5.6 we show the V_{exp}, N_H and τ_a distributions for the LAE population when $\langle X_{HI} \rangle = 0.3$ and $\phi_X = 0.0$. We find that the distributions are bimodal and that less than a 5% of the LAEs are outside the `FLaREON` accuracy regions (black dashed lines). On one hand, intrinsically bright LAEs have to decrease their number counts to fit the observed LF and they exhibit high V_{exp}, N_H and τ_a values. On the other hand, due to the IGM absorption the number density of LAEs struggle to fit observations. As discussed in previous chapters (see Fig. 2.3), the observed and intrinsic LAE LF are very close in the faint end at high redshift. As a Universe with $\langle X_{H_I} \rangle = 0.3$ has less volume where to place faint LAEs our model predicts that these galaxies exhibit high V_{exp} but low N_H and τ_a. In this way, our model tries to mitigate the decrease of faint LAE due to the IGM by boosting the ISM escape fraction.

After calibration, for each history of reionization we build our LAE samples. We select only galaxies with rest frame Lyα equivalent width EW > 20Å and $L_{Ly\alpha} > 3 \times 10^{42}$ erg/s. We have tested several Lyα luminosity cuts and we find the same results. Additionally, we have also tested different number density cuts

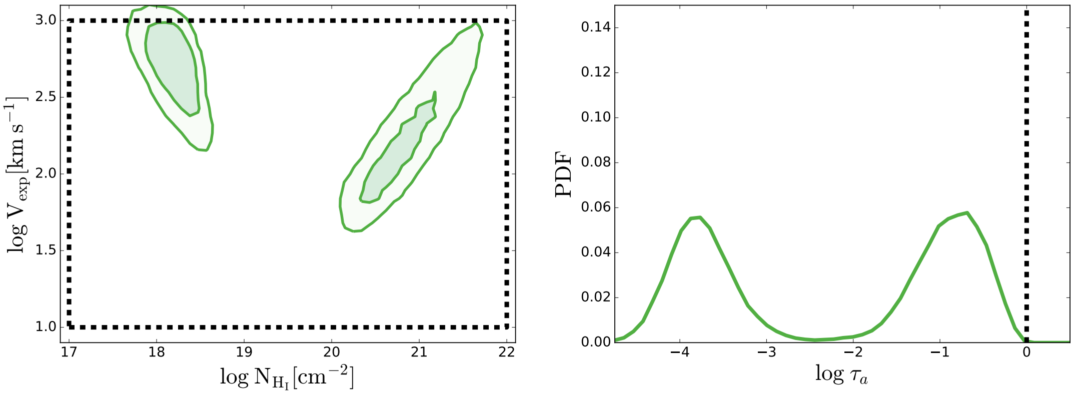


Fig. 5.6 Outflow properties of the LAE population for $\langle X_{H_I}(z = 6.7)\rangle = 0.3$ and $\phi_X = 0$. Left: V$_{exp}$-N$_H$ space. The dark and light regions show the 1σ and 2σ contours of the distribution. Right: τ_a distribution. In both panels the black dashed line indicates the high accuracy region of `FLaREON`.

ranking galaxies by observed Lyα flux and we obtained similar results.

In the previous chapter we placed the whole simulation at the same redshift. This allowed us to compute the IGM transmission at the edge (LoS$\sim$ 0) of the simulation using that the simulation was periodic. However, here, this condition no longer holds when we include the redshift dependence on the line of sight. This causes that there is a small fraction of the volume $z \sim 5.8$ which has a very short line of sight inside the simulation. In fact, we normally compute the IGM transmission up to ~ 35 cMpc h^{-1}, where the Lyα line is already strongly redshifted by the Hubble flow. To be safe we select only galaxies with coordinate along the line of sight greater than 42.16 cMpc h^{-1}, which reduces the simulation size along the line of sight to 500 cMpc h^{-1}. Therefore we also cut the simulation along the other two spatial directions. The final size of the volume populated by LAE is (500 cMpc $h^{-1})^3$. This corresponds to a 78% of the original volume of the simulation.

5.3 Results

In this section we analyze the main results of the LAE models during the epoch of reionization.

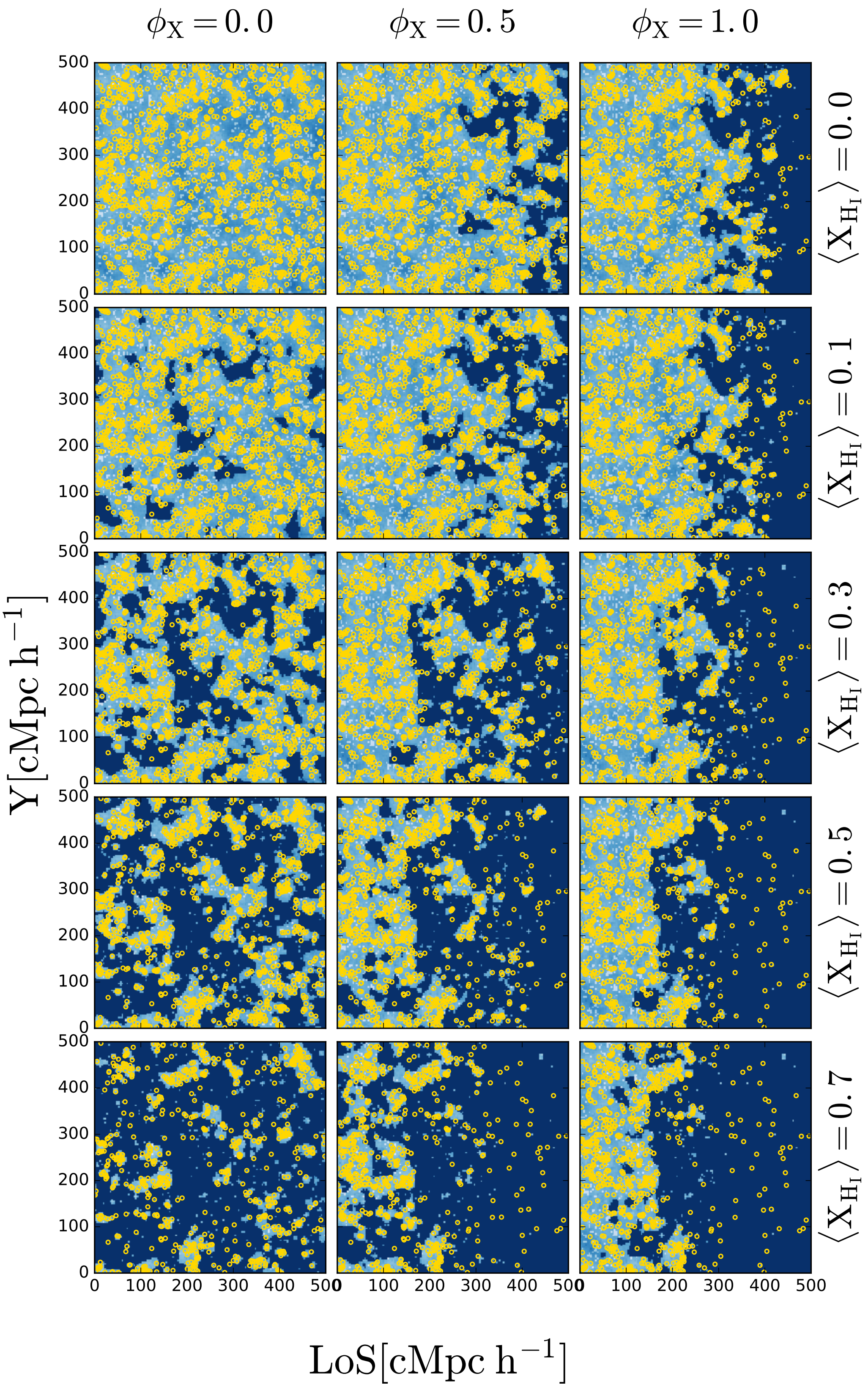


Fig. 5.7 Examples of spatial distribution of LAEs in comparison with the neutral hydrogen topology for different models (as labeled) in a slice of 3.6cMpc h^{-1}. In yellow rings we show observed LAEs. In dark blue we show neutral regions, while in light blue we show ionized bubbles.

5.3.1 LAE spatial distribution

In our model the history of reionization determines the $\Gamma(\overrightarrow{x})$ topology, which has a direct impact in $\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x})$. Additionally, the IGM transmission depends strongly on whether the galaxy is placed in a $\mathrm{H_{II}}$ bubble or a $\mathrm{H_I}$ clump. In the case in which the galaxy lies in a $\mathrm{H_{II}}$ regions the transmission redder to Lyα is 1, while it becomes lower than 1 if the galaxy is located in a $\mathrm{H_I}$ region. This causes that LAEs are preferentially observed in $\mathrm{H_{II}}$ bubbles.

In Fig. 5.7 we show the spatial distribution of LAEs for different reionization histories. In general, we find that LAEs (yellow rings) inhabit mainly in $\mathrm{H_{II}}$ regions (light blue areas) while they avoid $\mathrm{H_I}$ clumps (dark blue areas). Indeed, we find that the LAEs, in first order, follow the topology of $\mathrm{X}_{\mathrm{H_I}}(\overrightarrow{x})$. Additionally, the number of galaxies detected as LAEs depends on the history of reionization. For a reionization history with $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7) \rangle = 0.0$ and $\phi_X = 0.0$ the LAE distribution is relatively homogeneous in the whole volume. However, as a bigger fraction of the volume is covered by $\mathrm{H_I}$ the number of LAEs decreases and their spatial distribution becomes very inhomogeneous.

There is a small fraction of galaxies observed as LAEs that lie in $\mathrm{H_I}$ regions. These are galaxies that inject such a great flux of Lyα photons into the IGM that despite the large IGM attenuation, they are still detected as LAEs.

5.3.2 Number density of LAEs

In this section we study how the number density of LAE with Lyα luminosity greater than 3×10^{42} erg s^{-1} evolves along the line of sight for different histories of reionization.

In Fig. 5.8 we show the number density of LAEs along the line of sight $n_{\mathrm{LAE}}(\mathrm{LoS})$. In general, $n_{\mathrm{LAE}}(\mathrm{LoS})$ depends on the history of reionization. The number density along the LoS for the model with $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7) \rangle = 0.0$ and $\phi_X = 0.0$ represents the population with the greatest abundance of LAEs at any redshift. All the other histories of reionization can only remove LAEs from this sample.

In the models with $\phi_X = 0.0$ (purple lines), $n_{\mathrm{LAE}}(\mathrm{LoS})$ remains almost flat. However, there are more LAEs detected at lower redshifts. This would be caused by the

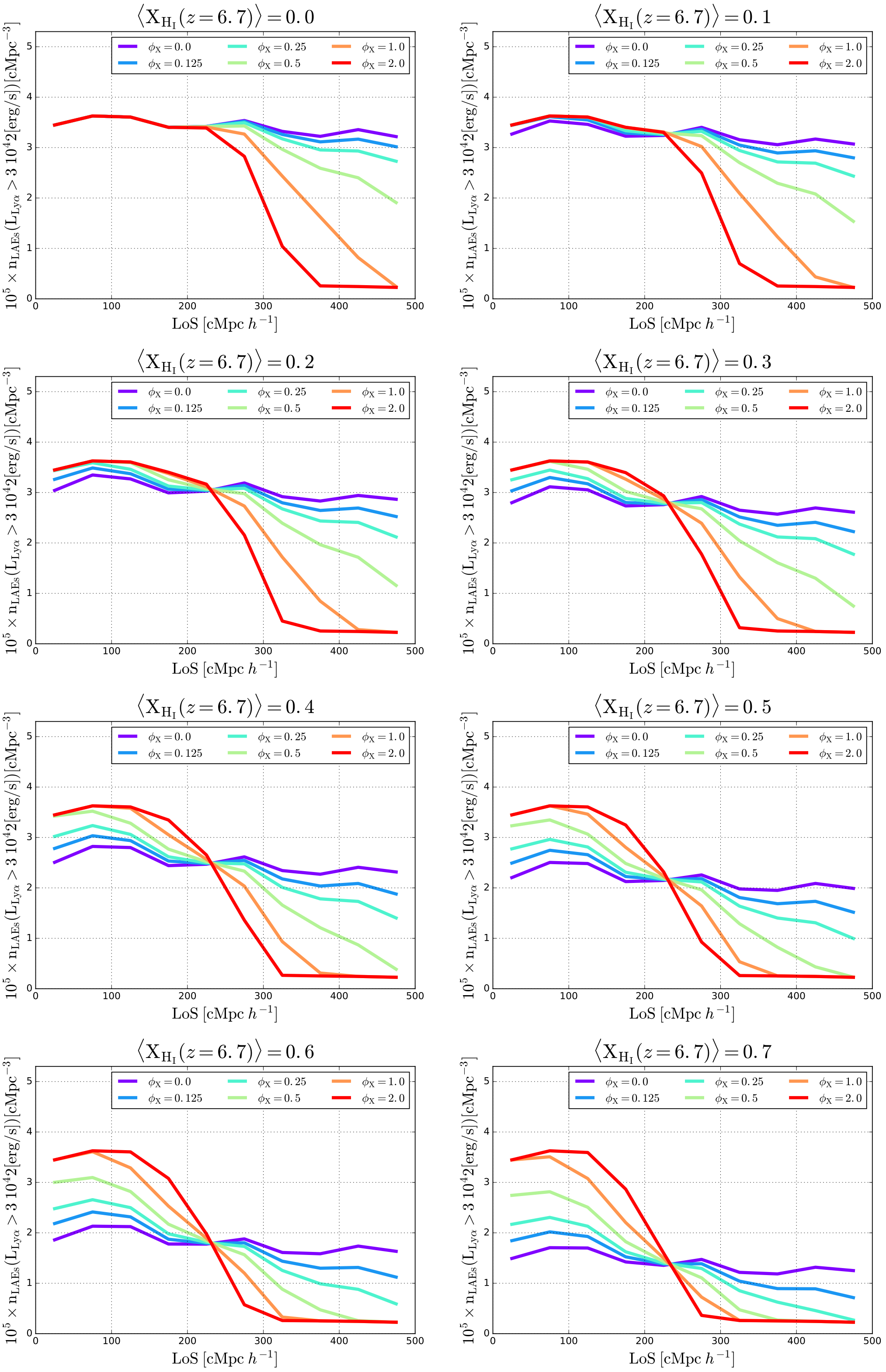


Fig. 5.8 Number density of LAEs along the line of sight for all the reionization histories. Each panel shows a fixed global fraction of neutral faction.

dependence of the IGM density on redshift. Indeed, at high redshift the IGM would be denser, thus more opaque to Lyα photons and less galaxies would be detected as LAEs. Additionally, increasing $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$ while keeping $\phi_X = 0.0$ causes a LoS independent decrease in $n_{\mathrm{LAE}}(\mathrm{LoS})$, as a greater fraction of the volume is covered by $\mathrm{H_I}$.

When $\phi_X \neq 0.0$ the global fraction of neutral hydrogen evolves across the volume, being higher at larger LoS distances. Indeed, in these models, $n_{\mathrm{LAE}}(\mathrm{LoS})$ depends on the distance. The higher the redshift is, the lower $n_{\mathrm{LAE}}(\mathrm{LoS})$ becomes. In particular, the higher ϕ_X is, the steeper is the evolution of $n_{\mathrm{LAE}}(\mathrm{LoS})$. Additionally, for a fixed value of $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$, the number density of all the models with different ϕ_X intersect in the same point. This behaviour is very similar to the one shown in Fig. 5.3. The intersection corresponds with $z = 6.7$, where all the $\langle \mathrm{X}_{\mathrm{H_I}}\rangle$ match. The number density value where these model intersect decreases with increasing $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$, as a smaller fraction of the volume is covered by $\mathrm{H_{II}}$.

For the half of the volume at highest redshift, at a given $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$, the number density of LAEs decreases as ϕ_X increases. Meanwhile in the half of the volume at the lowest redshift the opposite trend is found. This causes that the total number of LAEs in the whole volume depends on the combination of $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$ and ϕ_X.

5.3.3 Luminosity function

In Fig. 5.9 we show the LAE luminosity function of our models for fixed values of ϕ_{X}. Fig. 5.10 shows the different LAE LF for fixed values of $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$.

There is a clear dependence of the amplitude of the LAE LF with $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$. In particular, we find that, for a fixed ϕ_{X}, the amplitude of the LAE LF decreases as $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$ increases. However, this behaviour depends on the rate at which the Universe gets ionized. The amplitude variation are smaller the higher ϕ_{X}. As seen in the previous section, this is caused by the decrease in the ionized volume when augmenting $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$ for a fixed ϕ_{X}.

For a fixed value of $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$ we find that the amplitude of the LAE LF also depends on the rate at which the Universe gets ionized. At low $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$ the total number of LAEs anti-correlates with ϕ_{X}. However, for the highest value of $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$ considered here, this trend reverses. In these cases, the amplitude of the LF is higher for higher values of ϕ_{X}. In particular, in the extreme case where

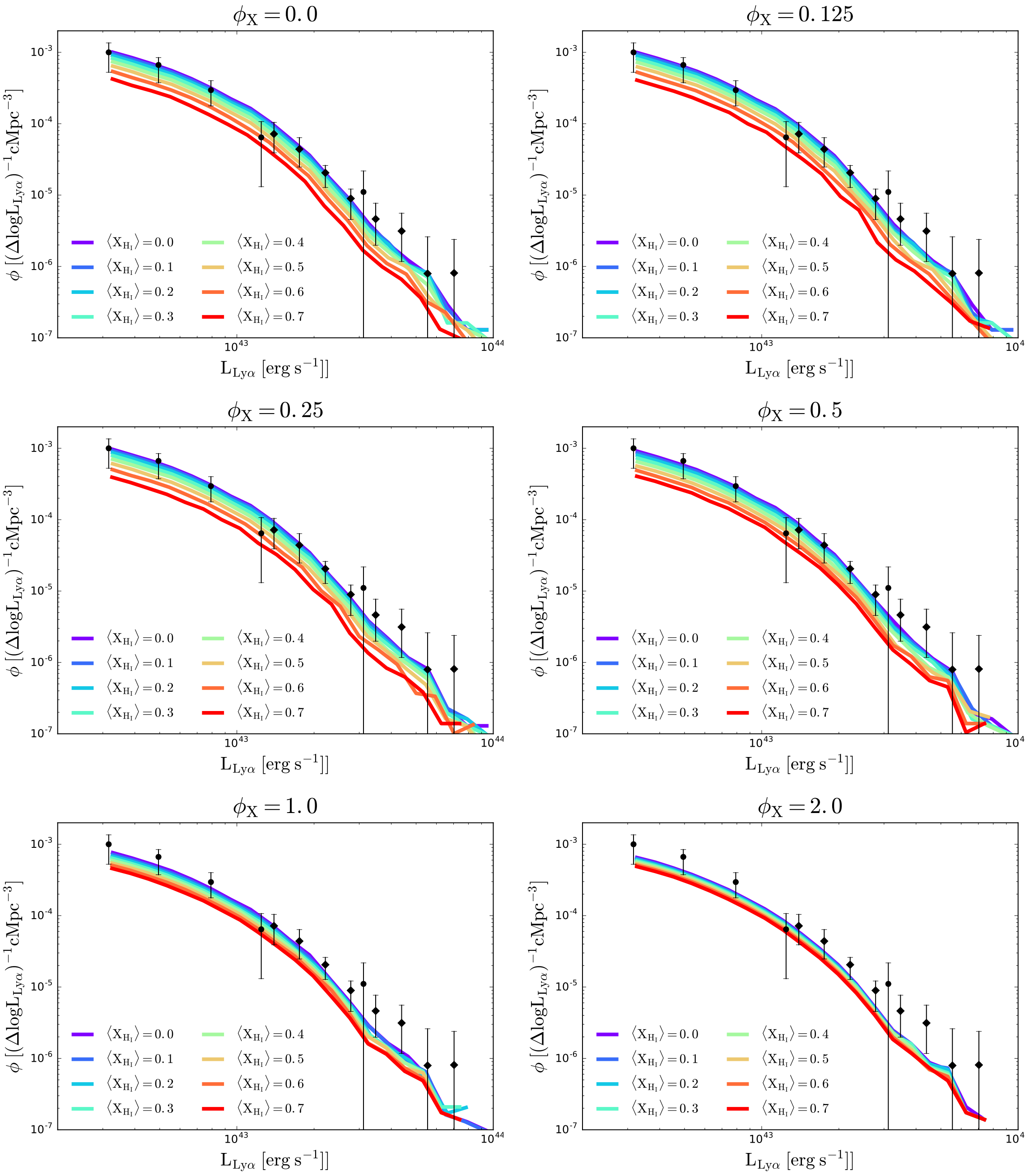


Fig. 5.9 LAE luminosity function for all the reionization histories. Each panel shows a fixed value of ϕ_X while $\langle X_{H_I}(z = 6.7) \rangle$ changes as labeled.

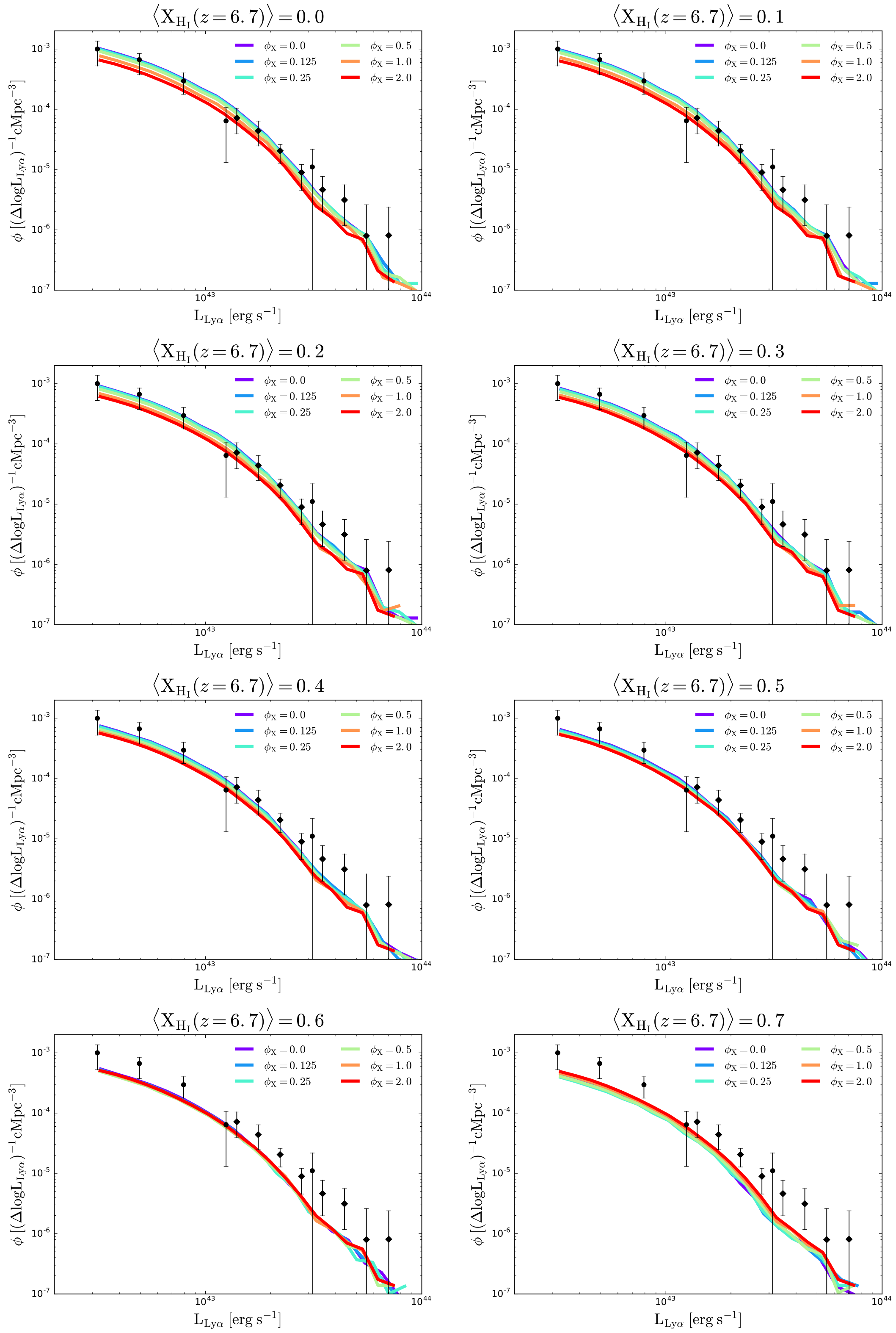


Fig. 5.10 Same as Fig. 5.9 but for fixed values of $\langle X_{H_I}(z = 6.7)\rangle$.

$\langle X_{H_I}(z=6.7)\rangle = 0.7$ and $\phi_X = 2.0$ the amount of H_{II} volume is greater than in the case of the same $\langle X_{H_I}(z=6.7)\rangle$ and $\phi_X = 0.0$, which translates into the different LF amplitudes.

5.3.4 Spherically averaged 2-point correlation function

In this section we study how the observed 3D clustering of LAEs depends on the history of reionization. For this, we compute the 3D 2-point correlation function (2PCF) of our models. In particular, we use the Landy-Szalay estimator (Landy and Szalay, 1993), i.e.,

$$\xi(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)} \tag{5.4}$$

where $DD(r)$ is the number of pairs of LAE-LAE at a given separation r. $DR(r)$ is the number of pairs of LAE and random particles uniformly distributed in the volume. Finally, $RR(r)$ is the number of random-random pairs. We set the integrated number density of random particles to be 10 times the integrated number density of the LAE sample with $\langle X_{H_I}(z=6.7)\rangle = 0.0$ and $\phi_X = 0.0$.

In Fig. 5.11 we show the 2PCF in real space of all the reionization histories used in this work. Overall, we find that the amplitude of the clustering increases with $\langle X_{H_I}(z=6.7)\rangle$. The change in amplitude also depends on the reionization rate. For small values of ϕ_X the changes in clustering amplitude are greater. As ϕ_X increases, the differences between different $\langle X_{H_I}(z=6.7)\rangle$ is smaller. This is caused by the similarity on the topologies of $\Gamma(\overrightarrow{x})$ between models with high ϕ_X, which translates in a similar spatial distribution of LAEs. Therefore, in similar clustering.

For fixed values of ϕ_X not only it changes the amplitude of the clustering, but also the shape of the 2PCF. The higher $\langle X_{H_I}(z=6.7)\rangle$ is, the greater the modification in the shape is. This behaviour also depends on ϕ_X. For example, in models with $\phi_X = 0.0$ the 2PCF exhibit a bump around ~ 25 cMpc h^{-1}. This bump is more bigger the bigger is $\langle X_{H_I}(z=6.7)\rangle$. This trend is the bump is not found for $\phi_X > 0.25$.

In Fig. 5.12, where we show the 2PCF in real space for fixed values of $\langle X_{H_I}(z=6.7)\rangle$. In general we find that the shape and the amplitude of the 2PCF also depend on ϕ_X. In particular, As we increase ϕ_X the 2PCF flattens at large scales. Additionally, at low

$\langle X_{H_I}(z = 6.7)\rangle$ the clustering amplitude of the clustering increases with ϕ_X. However, this trend progressively reverses towards higher $\langle X_{H_I}(z = 6.7)\rangle$.

5.3.5 Clustering parallel and perpendicular to the line of sight

In this section we study the clustering of LAEs along and perpendicular to the line of sight and how different reionization histories modify them.

In Fig. 5.13 we show the clustering of our LAE samples in real space divided in parallel and perpendicular to the LoS components. Over all, the clustering amplitude depends on both, $\langle X_{H_I}(z = 6.7)\rangle$ and ϕ_X. In particular, the clustering increases towards higher values of $\langle X_{H_I}(z = 6.7)\rangle$ and ϕ_X, as explored in the previous section.

We also find that the clustering amplitude long the parallel and perpendicular directions along the LoS behave different. For example, when $\langle X_{H_I}(z = 6.7)\rangle = 0.0$ and $\phi_X = 0.0$, both components are symmetric, as they were at lower redshifts (see Fig. 4.18).

However, as we increase ϕ_X the clustering parallel to the line of sight grows faster than the perpendicular one. This trends keeps up to $\phi_X \sim 0.5$, and from there on, the clustering perpendicular to the LoS is the strongest components. In contrast we find that increasing $\langle X_{H_I}(z = 6.7)\rangle$ makes the clustering parallel component grow faster for all the $\langle X_{H_I}(z = 6.7)\rangle$ explored in this work. Finally, the ratio between the clustering components is scale-dependent. This can be seen, for example, at $\langle X_{H_I}(z = 6.7)\rangle = 0.7$ and $\phi_X = 1.$, where on scales smaller than ~ 20 cMpc h^{-1} the clustering is symmetric, while on larger scales the perpendicular components dominates.

The behaviours found in real space have direct impact in the clustering in redshift space. In Fig. 5.14 we show the LAE clustering in redshift space divided in parallel and perpendicular to the line of sight. In this case, the amplitude of both components also increases with increasing $\langle X_{H_I}(z = 6.7)\rangle$ and ϕ_X. However, the shape has also the contribution of the redshift space distortions (Kaiser, 1987). Indeed, while in real space the LAE clustering in our model with $\langle X_{H_I}(z = 6.7)\rangle = 0.0$ and $\phi_X = 0.0$ was symmetric, in redshift space the perpendicular component is boosted. Additionally, in some cases (such as $\langle X_{H_I}(z = 6.7)\rangle = 0.0$ and $\phi_X = 0.5$) the redshift space distortions compensate the excess of clustering power parallel to the line of sight produced by the IGM, making the clustering symmetric.

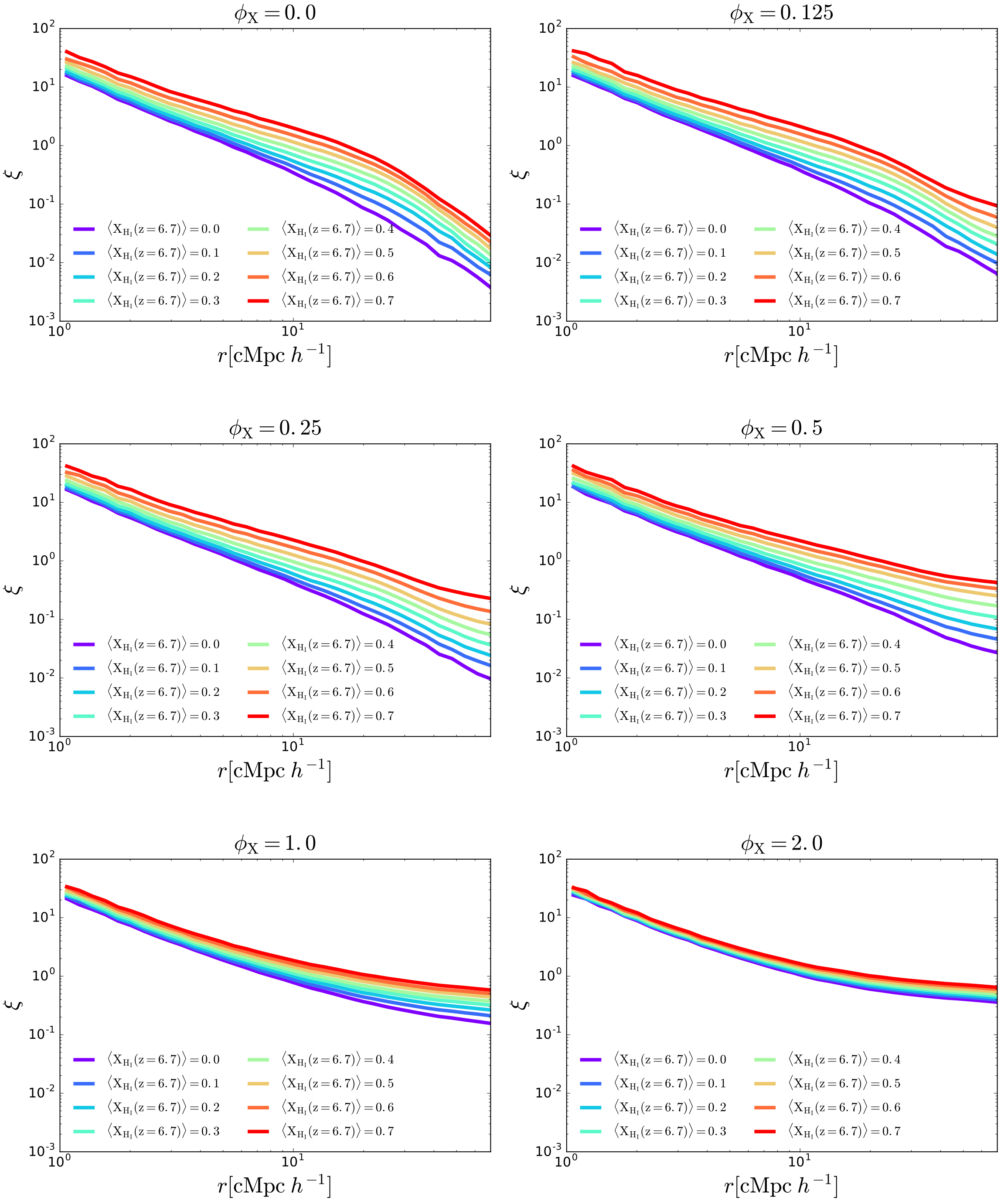


Fig. 5.11 LAE 2-Point correlation function for all the reionization histories. Each panel keep ϕ_X fixed, while $\langle X_{H_I}(z = 6.7)\rangle$ varies as labeled.

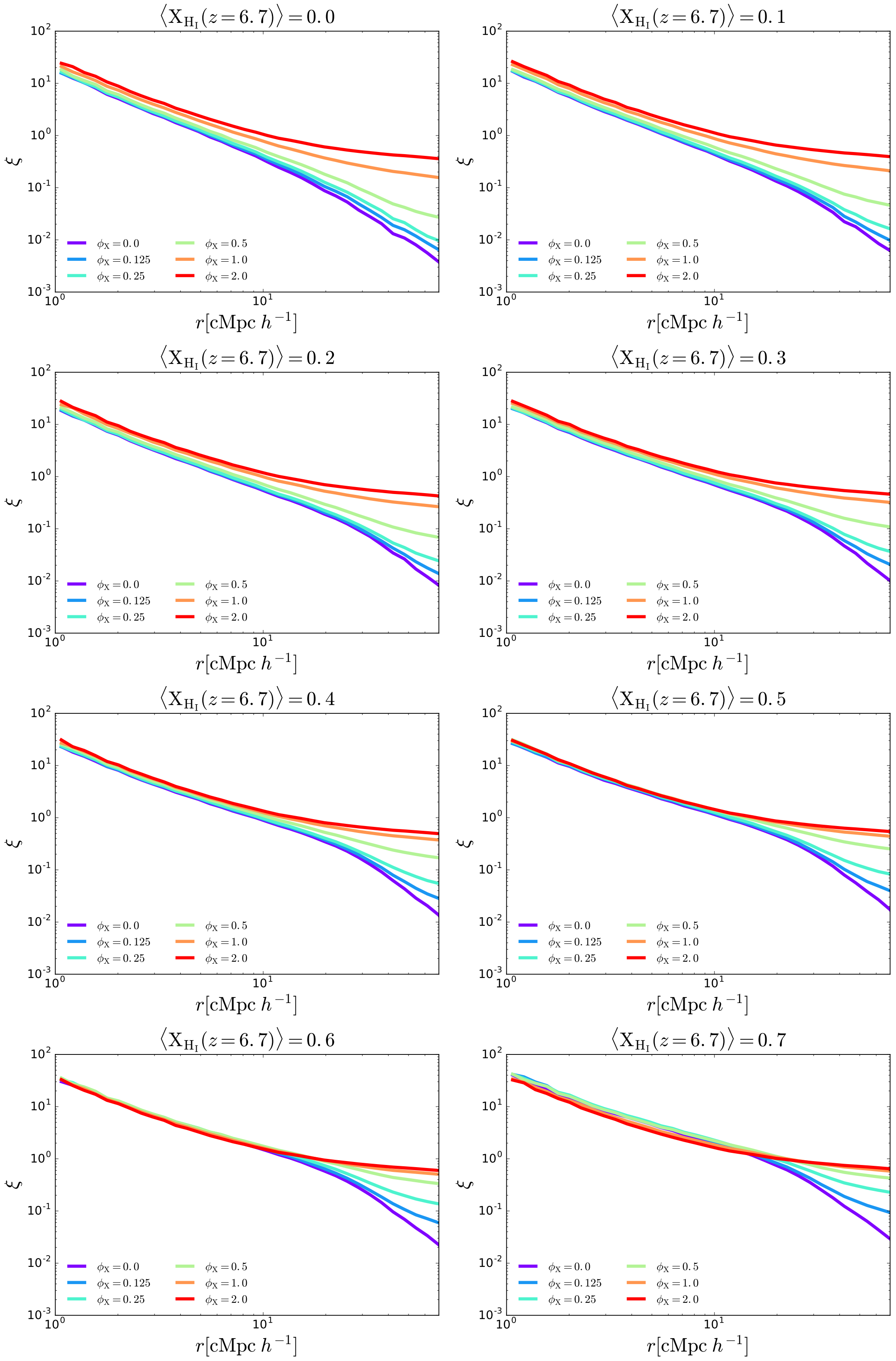


Fig. 5.12 Same as Fig. 5.11 but for fixed values of $\langle X_{H_I}(z = 6.7)\rangle$.

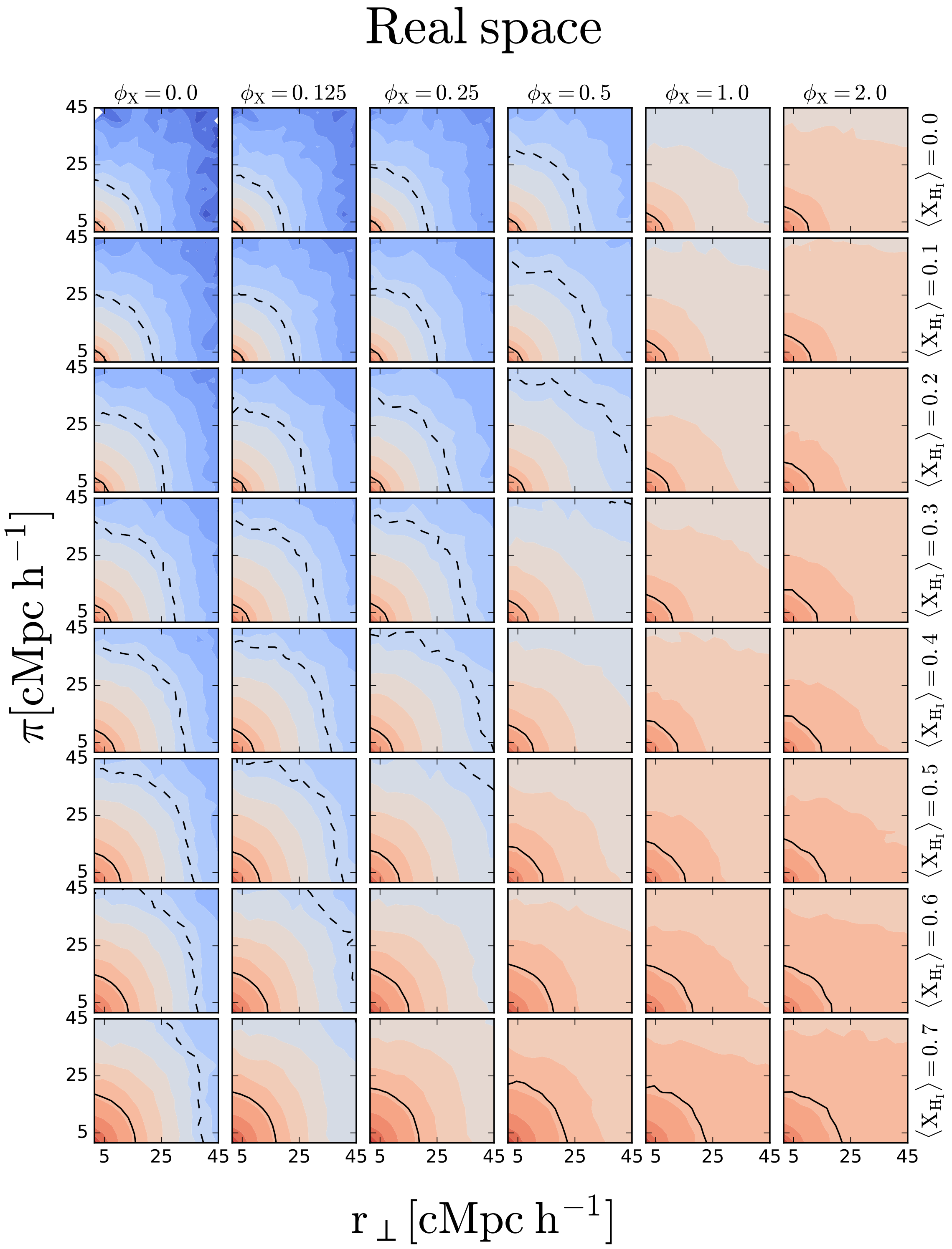


Fig. 5.13 Clustering divided in parallel and perpendicular to the line of sight components in real space. The dashed and solid black lines corresponds with $\xi(\pi, r_\perp) = 0.1$ and 1.0 respectively. Columns have a fixed value of ϕ_X and rows a fixed value of global fraction of neutral hydrogen at redshift 6.7 as labeled.

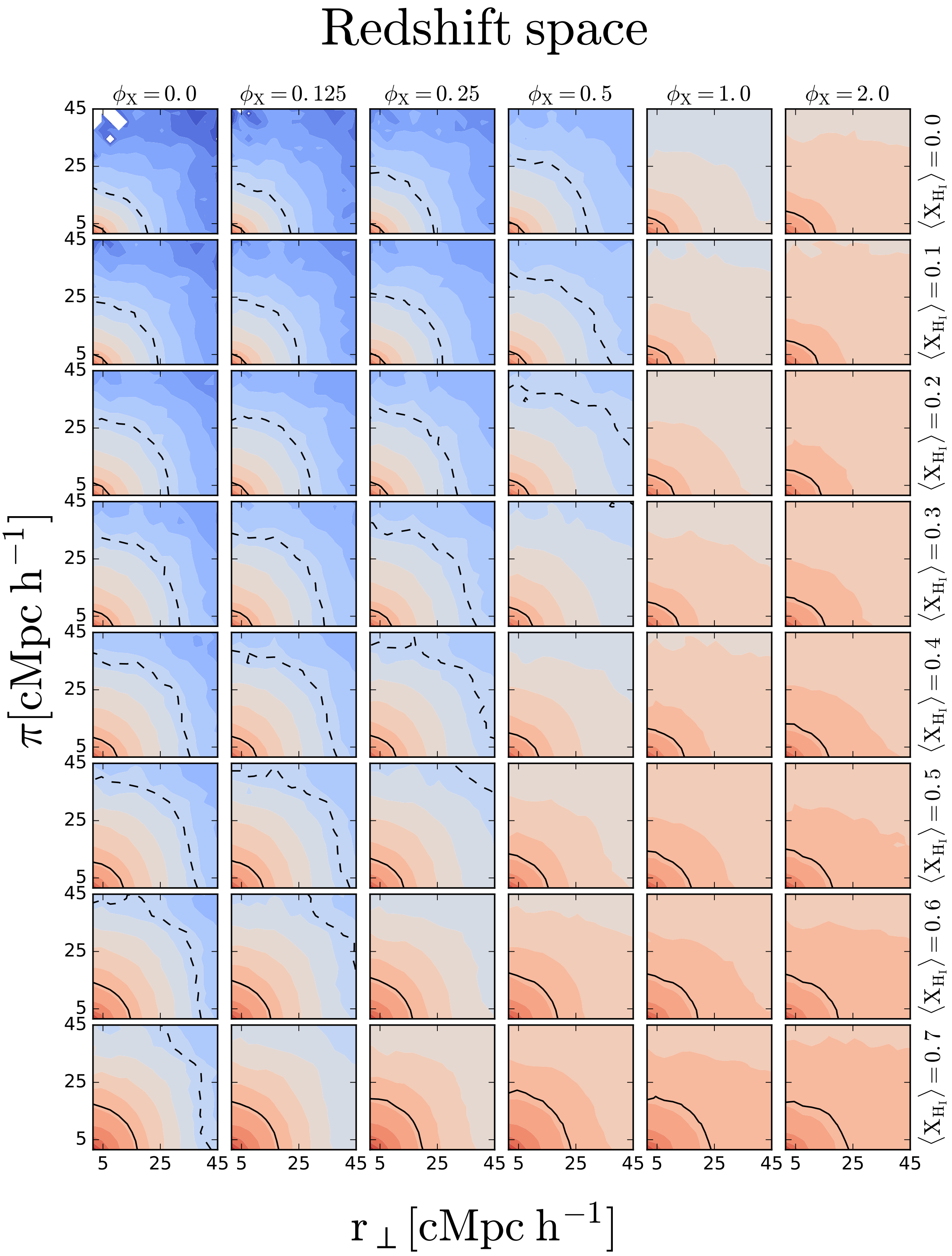


Fig. 5.14 Same as Fig. 5.13 but in redshift space.

5.3.6 The clustering quadrupole of LAEs

In this section we analyze the quadrupole of our different LAE samples in real and redshift space. We compute it by Eq. 4.31.

In Fig. 5.15 we show the LAE quadrupole in real and redshift space for fixed values of ϕ_{X}. Over all, we find that the shape and amplitude of the quadrupole depends on the history of reionzation. In our model with $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle = 0.0$ and $\phi_{\mathrm{X}} = 0.0$ we find that the quadrupole is very close to 0 in real space. However, increasing $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)$ for low values of ϕ_{X} creates a valley around ~ 25 cMpc h^{-1}. For greater values of ϕ_{X}, increasing $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)$ results into a big increase in the quadrupole amplitude on scales larger than ~ 35 cMpc h^{-1}. Moreover, the higher the global fraction of neutral hydrogen is at the center of our box, the higher is the boost that experiments the quadrupole. Additionally, we find that in real space the quadrupole is close to 0 for distances close to 0.

In redshift space the quadrupole is modified and the amplitude of the quadrupole increases on scales smaller than ~ 15 cMpc h^{-1}. Then, the redshift space distortions do not contribute significantly on larger scales.

In Fig. 5.16 we show the quadrupole for different reionization histories in real and redshift space for fixed values of $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)$. The LAE samples with $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle = 0.0$ and low values of ϕ_{X} exhibit a quadrupole flat a comparable with zero. However, even if $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle = 0.0$, for high values of ϕ_{X} the quadrupole exhibits a boot on large scales. Additionally, we find that for a fixed $\langle \mathrm{X}_{\mathrm{H_I}}(z = 6.7)\rangle$, increasing ϕ_{X} always increases the amplitude of the boost. Finally, the same trends are found in redshift space.

5.4 Conclusions and future work

In this chapter we have studied how different histories of reionization change the properties of Lyα selected samples. For this goal we have implemented a reionization scenario in the LAE model presented in Chapter 4.

In order to mimic the typical topology of the local fraction of neutral hydrogen during the epoch of reionization we have introduced a method by which the regions

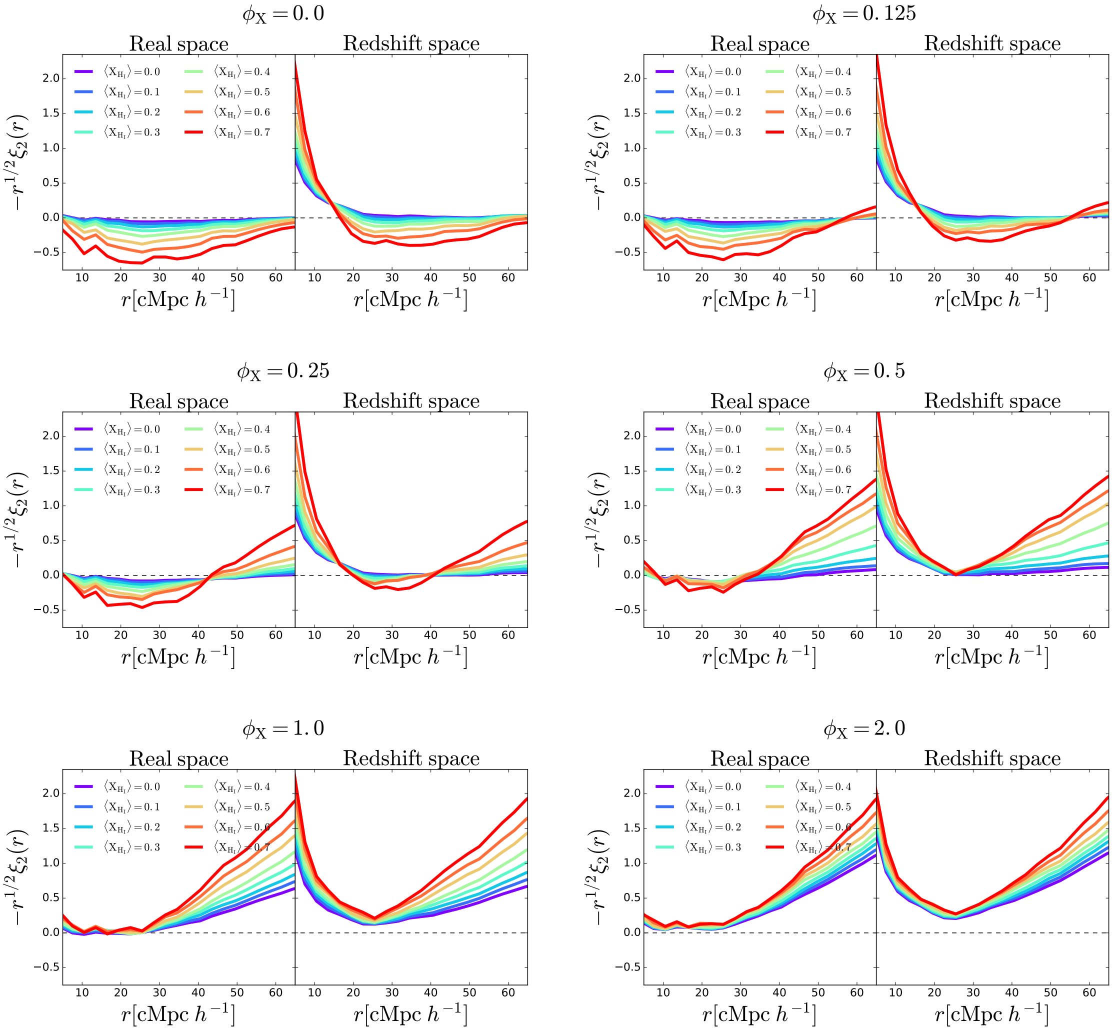


Fig. 5.15 Quadrupole of LAEs for all the reionization histories studied in this chapter in real and redshift space. Each panel has a different fixed value of ϕ_X while $\langle X_{H_I}(z = 6.7)\rangle$ varies as labeled.

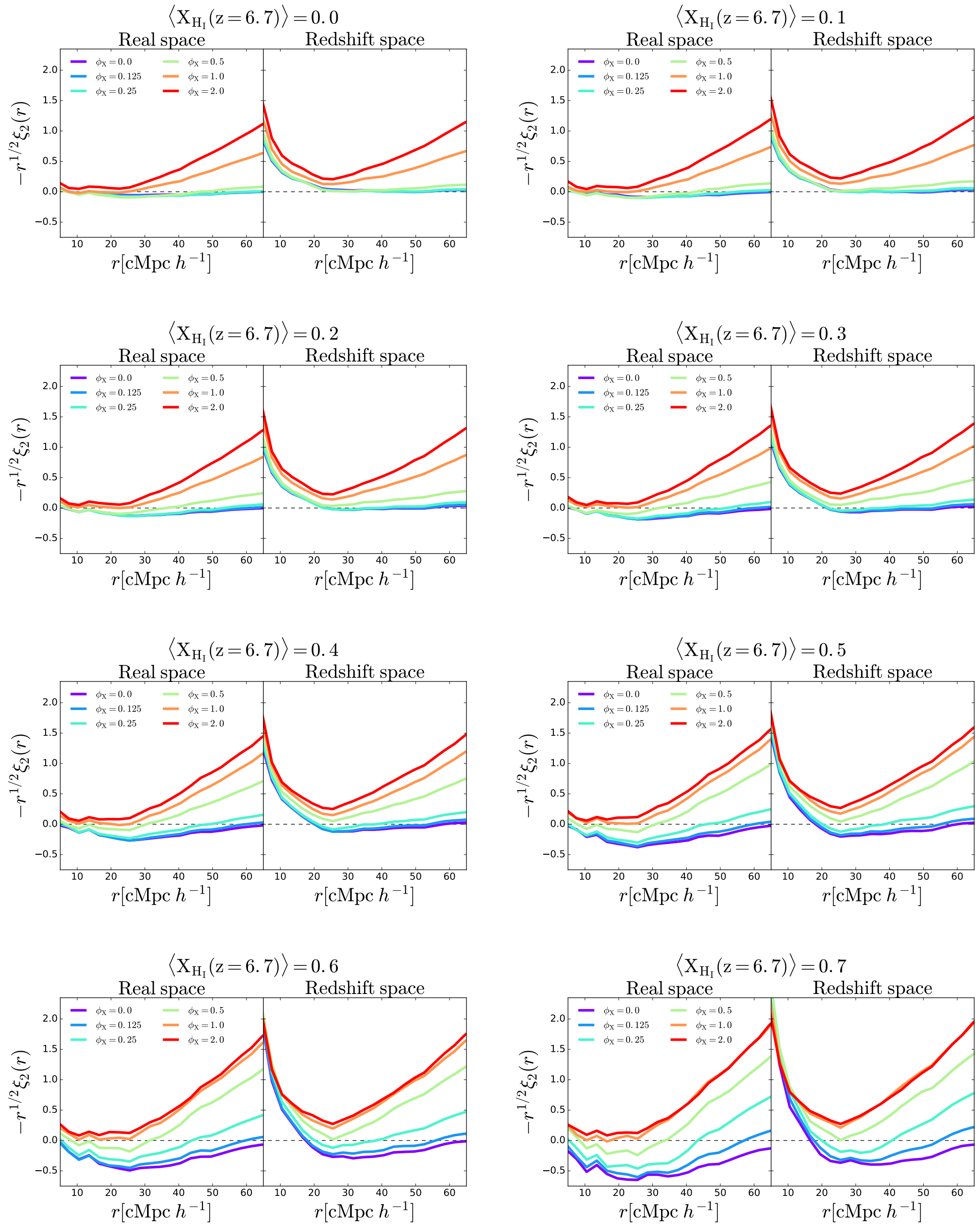


Fig. 5.16 Same as Fig. 5.15 but for fixed values of $\langle X_{H_I}(z = 6.7)\rangle$.

with the lowest photoionization rate are considered shelf-shielded. This divides the simulation volume in two kind of regions: $\mathrm{H_I}$ clumps and $\mathrm{H_{II}}$ bubbles. Inside the first ones almost all the hydrogen is neutral, while in the second ones $\langle \mathrm{X_{H_I}} \rangle \sim 2 \times 10^{-4}$. The $\mathrm{H_I}$ clumps have lower transmission around Lyα than the $\mathrm{H_{II}}$ bubbles. This causes that galaxies lying in $\mathrm{H_{II}}$ regions are preferentially observed as LAEs.

Then, we have made a simple toy-model for producing different histories of reionization. Our toy-model is based on the global neutral hydrogen fraction at the center of our simulation ($\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$) and a parameter that controls how fast reionization is completed (ϕ_X). We have sampled theses parameters to create several histories of reionization, generating a total of 48 $\langle \mathrm{X_{H_I}}(z) \rangle$ evolutions. We have computed the IGM transmission for each position of our simulation for each of them. Then we have run our LAE model for each reionization history.

Consecutively we have studied how the different histories of reionization affect the properties of LAEs. For example, the LAE luminosity function changes its amplitude with $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$ and ϕ_X. We find that the LF amplitude always correlates with $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$, while the dependence with ϕ_X is not that clear. However, the shape is maintained across different histories. In this way, the effects of changing $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$ and ϕ_X are degenerated, which makes the LAE LF a poor tool to estimate the history of reionization.

A promising tool to unravel the history of reionization is the clustering of LAEs. We find that changing the reionization rate and its time changes differently the 2-point correlation function of LAEs. In particular, the clustering amplitude correlated with $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$ and the shape on large scales with ϕ_X. Additionally, we find that when diving the clustering in parallel and perpendicular to the line of side, changing $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$ and ϕ_X impact in different ways too. These behaviours could help to constrain the history of reionization with future LAE surveys.

In a near future we plan to upgrade the model presented here. For building this model we have assumed several things that could be improved. We plan to build a light cone in our simulation from z=5.7 to z=7.9. This will improve the model in different ways. For example, the population of galaxies will evolve with distance to the observer (here we assumed it constant). Also, the comoving density field will change along the distance to the observer.

Chapter 6

HiRULE : High Redshift Universe Lyman-α survEy

Abstract

We present the first results of HiRULE: High Redshift Universe Lyman-α survEy, a program targeting LAEs and QSOs at $z \sim 2.2$ to study their cross-correlation over an unprecedented volume. In its initial phase, HiRULE consists in observing 20 deg^2 with the narrow band J0395 at T80Cam in the Observatiorio Astrofísico de Javalambre, allowing us to perform various cross-correlation studies of these two high-z populations. Key science cases include: a) Cross-correlation function of LAEs and QSOs, to quantify baryonic environmental effects acting at this redshift and affecting the small-scale clustering around QSOs; b) Lyα fluorescence around QSOs, from neutral gas surrounding QSOs; c) the properties of bright Ly-alpha emitters at $z \sim 2.2$, including their luminosity function and auto-correlation function. Here we present the data analysis and the measured LAE luminosity function.

6.1 Introduction and scientific rationale

Multiwavelength studies of high redshift galaxies are typically performed over small areas, exploiting large telescopes to reach faint, numerous objects. Pioneering campaigns attempting to study the nature of Lyman-alpha emitters at high redshifts are the LaLa survey (Malhotra and Rhoads, 2002), the MUSYC survey (Gawiser et al., 2006), SXDS (Ouchi et al., 2005) and HiZELS (Geach et al., 2008). These studies have revealed the nature of Lyα emitters (LAEs) as compact, young and typically dust-poor objects

(Gronwall et al., 2007; Guaita et al., 2010). Some groups have attempted to measure the clustering of this galaxy population with a few hundred objects at most, despite these measurements being evidently dominated by cosmic variance (Orsi et al., 2008).

The High Redshift Universe Lyman-alpha survey (HiRULE) is an initiative to obtain and characterize a big LAEs sample at redshift 2.2 using deep J0395 narrow band photometry reaching a high number density. Compared to previous narrow-band surveys, HiRULE is shallower, but covers a sky area that is between 1 and 2 orders of magnitude larger than any previous Lyα survey at this redshift, making it the largest photometric survey of LAEs. In addition, the completion of HiRULE, if executed in due time, should anticipate the results from HETDEX, an IFU (integral field unit) survey of LAEs at 1.9<z<3.8 over $\sim$ 300 deg^2 being currently carried out.

With such large area, the data can be used to explore science cases that were impractical in previous surveys due to their prohibitively small area. In particular, the cross-correlation with QSOs at the same redshift is a science topic that has not been explored at this redshift before. In the standard hierarchical galaxy formation scenario, QSOs are typically regarded as being hosted by dark matter haloes of masses around $10^{12} M_\odot/h$. Hence, it is typically expected that these extreme objects might be pinpointing the progenitors of massive galaxy clusters. Such scenario, however, depends on the assumed physical mechanisms driving the coevolution of super-massive black holes and their host galaxy. This is why a cross-correlation clustering study of QSOs at high redshifts should reveal the typical halo masses that these objects are hosted by, and any environmental baryonic mechanism that might already be in place at high redshits. This is the backbone of the motivation for HiRULE.

The strategy of HiRULE takes advantage of the unique features of the Observatiorio Astrofísico de Javalambre (OAJ) telescopes, namely the large FoV and narrow band photometry.

The scientific goal of this survey is to measure the cross correlation between LAEs and QSO at redshift 2.2. The big area and great depth will allow us to compute this cross-correlation for the first time ever. Previous attempts to cross-correlate QSO and high redshift galaxies have been made in significantly smaller areas (García-Vergara et al., 2017; Garcia-Vergara et al., 2019).

Table 6.1 Description of each of the different field composing the HiRULE survey. This table lists the area overlapping with HSC (Aihara et al., 2018), the depth, number of LAEs and median color in each of the field.

Field	ra [*deg*]	dec [*deg*]	Area [deg^2]	Depth (S/N=5)	N_{LAEs}	$\langle$g – J3950$\rangle$ [*mag*]
HiRULE-00	331.1962	-0.1934	0.2654	22.40	24	-0.84
HiRULE-01	331.1962	1.2002	0.7941	22.20	95	-0.63
HiRULE-02	331.1962	2.5933	0.1308	22.00	9	-0.64
HiRULE-10	332.5920	-0.1934	1.3797	21.90	81	-0.77
HiRULE-11	332.5920	1.2002	1.6909	21.95	180	-0.55
HiRULE-12	332.5920	2.5933	0.3070	21.83	28	-0.64
HiRULE-20	333.9878	-0.1934	1.3471	21.68	48	-0.69
HiRULE-21	333.9878	1.2002	1.6179	22.12	133	-0.68
HiRULE-22	333.9878	2.5933	0.2798	22.21	13	-0.85
HiRULE-30	335.3832	-0.1934	1.3258	22.39	96	-0.79
HiRULE (total)	-	-	9.3247	-	707	-

Lyα emission gets attenuated by the circumgalactic (CGM) and intergalactic medium (IGM). Additionally it is also thought that QSO might be one of the main sources of ionizing photons for z<4 so if this is the case the cross-correlation between these two galaxy population should be positive. In the opposite direction, as we studied in Chapter 4 the observability of Lyα is coupled to the IGM large scale properties. In the case of star forming galaxies emitting Lyα, their line profile is relative narrow and normally the FWHM of the lines is below 10Å. In comparison, QSOs exhibit much broader lines that extend redder than Lyα, reaching $\sim$ 100Å of FWHM. In this scenario star forming galaxies are more affected by the IGM than QSO, as only a very small fraction of the Lyα flux in QSO is actually close to Lyα. This makes that the coupling with the IGM is weaker in QSO. Star forming LAEs thus tend to populate regions with $\mathrm{V_z} > 0$ (see Chapter 4), while QSO are distributed without any selection effect on the IGM. This could decrease the cross-correlation between these populations. Garcia-Vergara et al. (2019) studied the cross-correlation between star forming LAEs and QSO and found that it was lower than expected.

Recombination of ionizing photons by the neutral CGM around a QSO, and subsequent Lyα scattering is expected to result in fluorescent emission. Such effect can be measured by stacking the photometry of QSOs, and this constitutes another key project within HiRULE.

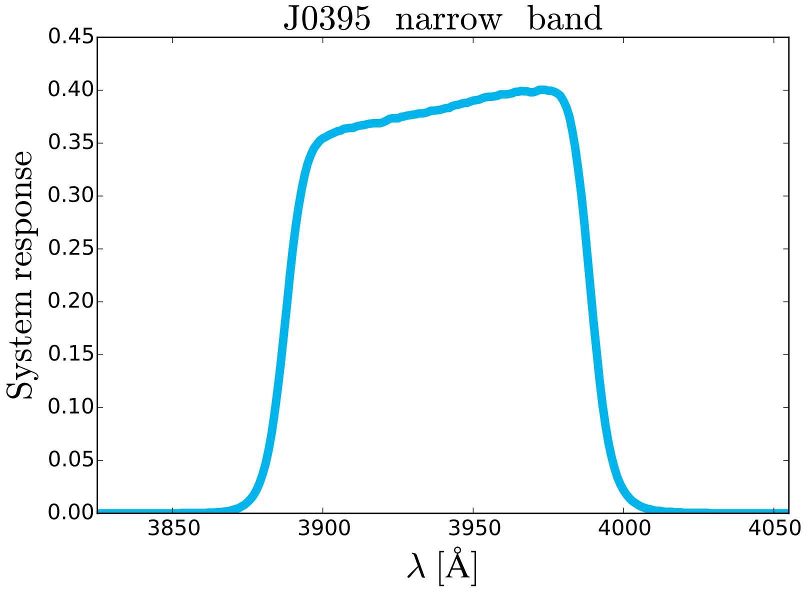


Fig. 6.1 System response of the narrow band J0395.

Other key scientific cases are

1. The auto-correlation function of LAEs at $z \sim 2.2$.
2. The bright end of the LAEs Luminosity function is a quite controversial at the moment. While some work detect a bump Konno et al. (2016) in the bright end in others is not present Sobral et al. (2017b). The HiRULE wide area can manage to solve this controversy.
3. As we will use also broad band photometry the EW distribution of LAEs could also be studied.

The narrow-band J0395 will provide LAEs between redshift 2.205 and 2.275. In order to determine the source redshift we will also use broad band Hyper Suprime Cam Subaru Strategic Program (HSC-SSP) photometry (Aihara et al., 2018).

In a future phase, we plan to increase the volume with adjacent redshift windows using the J-PAS narrow band filter set or even use HiRULE as a starting point to perform a larger and deeper study of LAEs with JPCAM@JST.

6.2 Technical description

The data acquisition took place in the second half of 2017 during several nights. The typical seeing during these nights was 1.2 *arcsec* and the sky quality was *dark*. A total of 10 different fields were observed with the narrow band J0395 at T80Cam in the OAJ. The system response, i.e., the convolution of the T80Cam response and the J0395 transmission curve, is shown in figure 6.1. The full width half maximum of the filter is 101Å and the filter is centered at 3950Å.

T80Cam has a square field of view of $\sim 1.4 \times 1.4\ deg^2$. Therefore, the total sky coverage of HiRULE is $20\,\mathrm{deg}^2$. The coordinates of the different field are listed in Table 6.1. Additionally, in Fig. 6.2 we display the footprint of HiRULE and the HSC overlapping region and the code name of each field. The fields are named as follows : 'HiRULE-' + 'a' + 'b', where a labels the position of the field in right ascension (RA) and b in declination (DEC). In this way, HiRULE-00 is the field with lowest RA and DEC. We also present a logarithmic render of the field HiRULE-30 at the end of this chapter (6.7).

Each of the 10 fields were observed in rounds of 6 exposures of 10 minutes. The individual exposures were slightly shifted, making a dithering pattern to eliminate spurious sources and background inhomogeneity. Then, using J-PLUS (Cenarro et al., 2019) pipeline they were combined into single co-added images from which the source catalogs were extracted. The depth, defined as the magnitude at which the detection reach a signal to noise ratio of 5, of each field is listed in Table 6.1.

We masked regions contaminated by bright stars and the edges of the images, where the noise is higher due to the dithering pattern. The mask in the overlapping regions of HSC and HiRULE is shown in Fig. 6.3. The blue areas correspond to regions where the quality of the photometry was good. From now on we will work only with the sources that lie in these sky patches.

The final sky coverage of each of the HiRULE fields overlapping HSC is very different (see Table 6.1). Additionally HiRULE does not cover the whole HSC footprint even if the HiRULE total sky coverage is sufficient for it. The original idea for HiRULE was to observed a total of 18 field and cover completely the HSC footprint. Sadly, due to the weather conditions of the second half of 2017 the acquisition of the full data set was unfeasible.

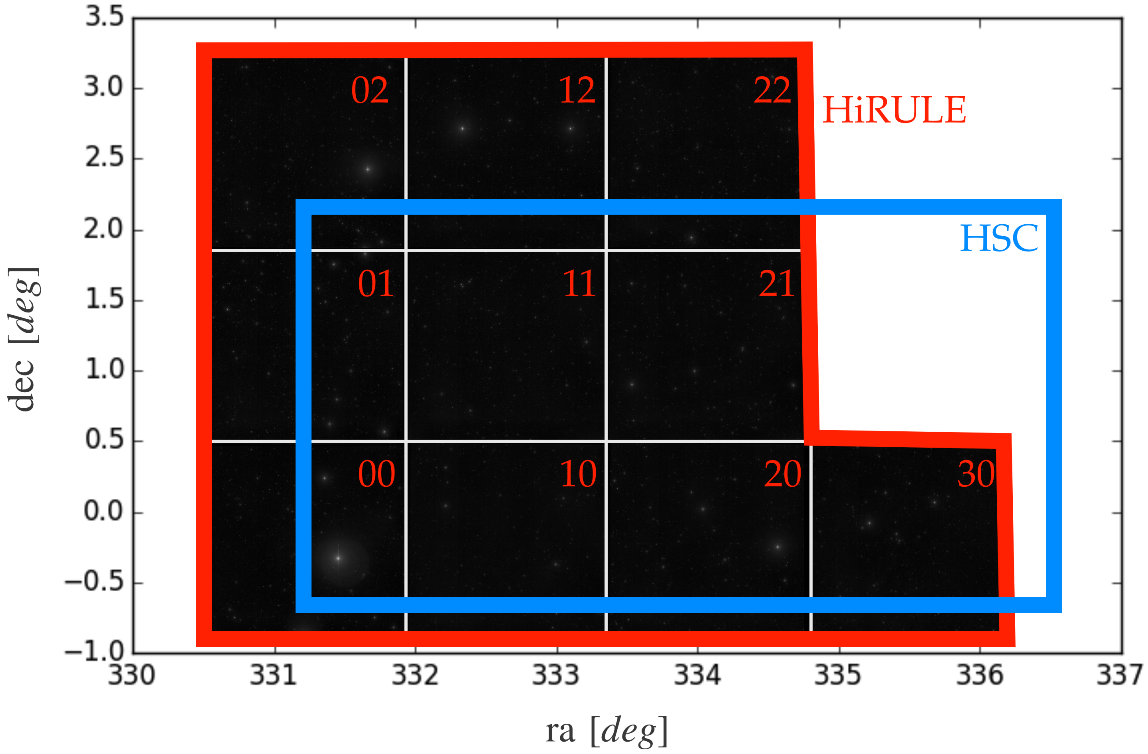


Fig. 6.2 Footprint of HiRULE (red) and the HSC overlapping field (blue). The red number indicate the code-names each HiRULE field.

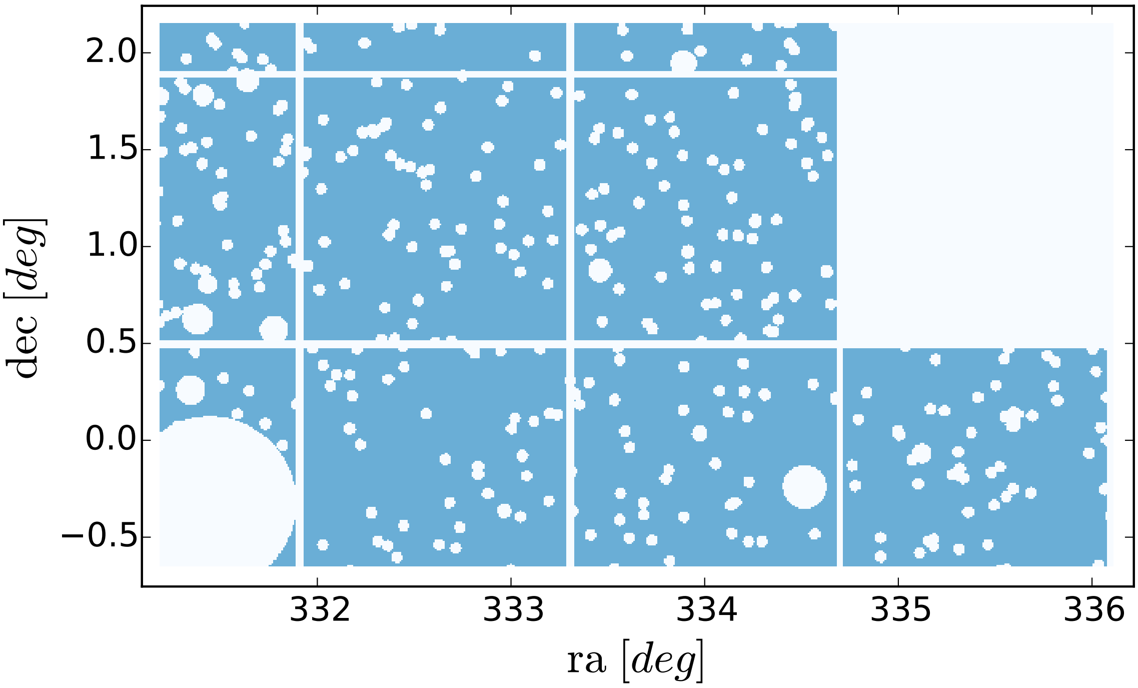


Fig. 6.3 Mask of the overlapping regions between HiRULE and HSC. In blue we show the high quality regions and in white the unused/unobserved ones.

6.3 LAE candidate selection

In this section we describe how the LAE candidate selection is made from the HiRULE observations.

HiRULE is composed by 10 different fields. The quality of the nights when these field were observed changed from night to night. Then, the data reduction was done field by field following J-PLUS pipe line (Cenarro et al., 2019). As a consequence each field has a different depth. Therefore, for the candidate selection we treat each field independently.

In short, we make a candidate selection based on colors. We make use of magnitudes measure in apertures of 3 *arcsec*. A source has to full fill all of the next four requirements in order to be consider a LAE candidate:

1. The signal to noise ratio of the source has to be greater than 5.

2. Their magnitude in J0395 has to be less than 19.5 to avoid stars and artifacts.

3. The color g-J0395> 0.19, where g is the broadband photometry from HSC (Aihara et al., 2018). This color cut translate in a threshold in rest frame equivalent width of 20Å

4. The color excess respect to the median color of the field has to exceed 3 times the 1σ error of the color. This is,

$$\Delta(\mathrm{g}-\mathrm{J3950}) = (\mathrm{g}-\mathrm{J3950}) - \langle \mathrm{g}-\mathrm{J3950} \rangle > 3 \times \sigma(\mathrm{g}-\mathrm{J3950}), \tag{6.1}$$

   where $\langle \mathrm{g}-\mathrm{J3950} \rangle$ is computed in $14 < J0395[mag] < 19.5$. As a result, $\langle \mathrm{g}-\mathrm{J3950} \rangle$ is slightly different in each field. Its values are listed in Table 6.1.

The LAE candidate selection is shown is Fig. 6.4. The position of each panel matches with the sky disposition oh HiRULE (see Fig. 6.2). In colors we show all the sources subtracted from the HiRULE images and in black those selected as LAE candidates. Over all, the cut in S/N=5 avoids the selection of spurious sources present at greater magnitudes with colors up to 5 magnitudes. These are faint detections with great uncertainty in HiRULE with a counter part in HSC with very low continuum. This combination makes that these objects apparently have a huge color excess. Additionally, in most of the cases the cut in equivalent with is more restrictive than the one in color. As a result, most of the sources are just selected by their equivalent width and S/N.

The number of LAE candidates in each field is listed in Table 6.1. The abundance of candidates varies from field to field. There are mainly two reasons for this: i) the sky area covered by each field is different and ii) the sky coverage of each field is different. As a result, field with lower depth and sky coverage exhibit a lower number of LAEs (as HiRULE-02).

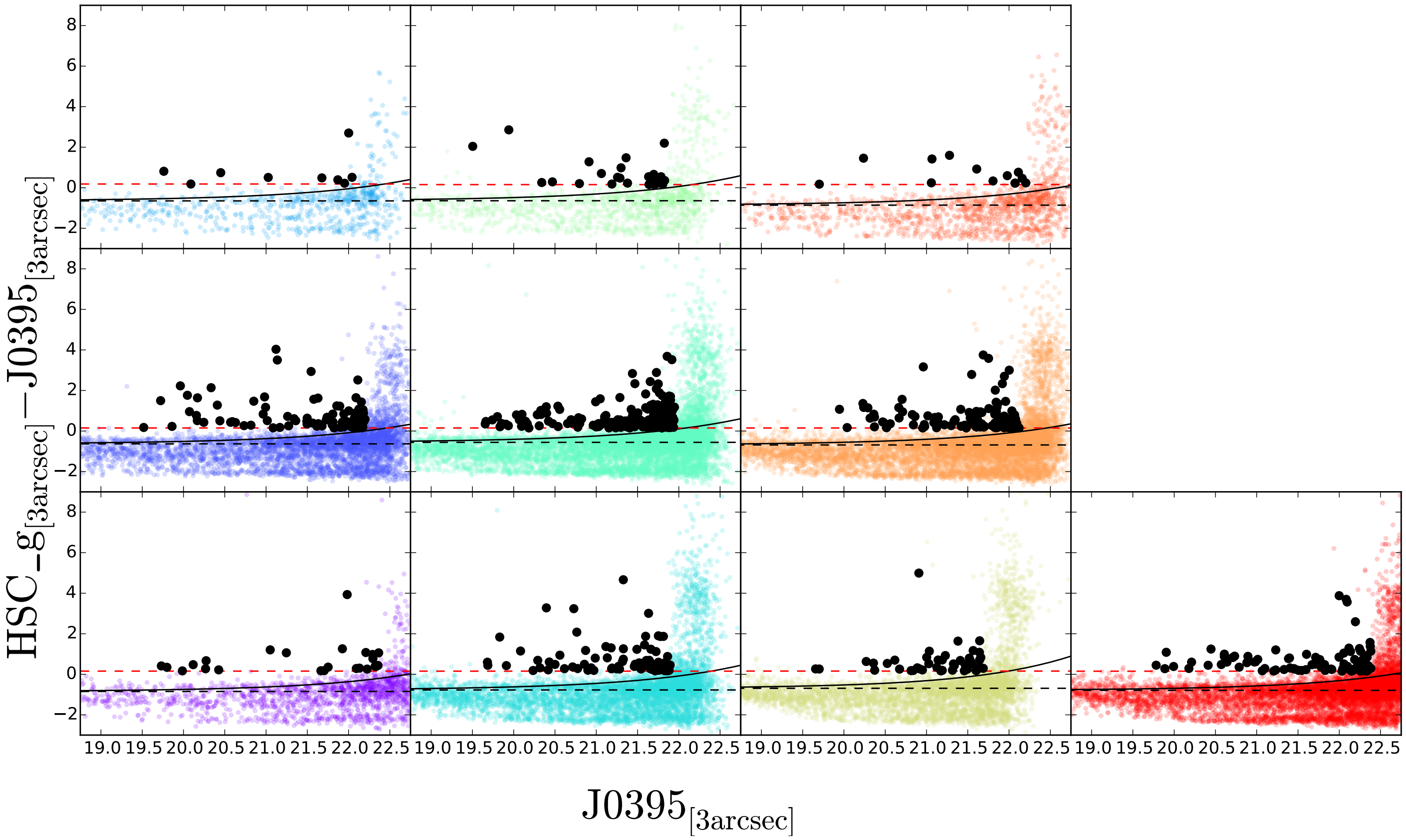


Fig. 6.4 Color selection of LAE candidates. Each panel is a different HiRULE field. The disposition is the same as in the sky (see Fig. 6.2). In this way purple shows HiRULE-00, red shows HiRULE-30 and so on. In colored dots we show all the sources detected in each field. Black dots are sources selected as LAE candidates. The dashed black line shows the mean color in each field. In dashed red we show the color cut corresponding to a rest frame equivalent width of 20Å. The solid black line shows the 3σ error with respect the median color. This is different in each field since the depth is different.

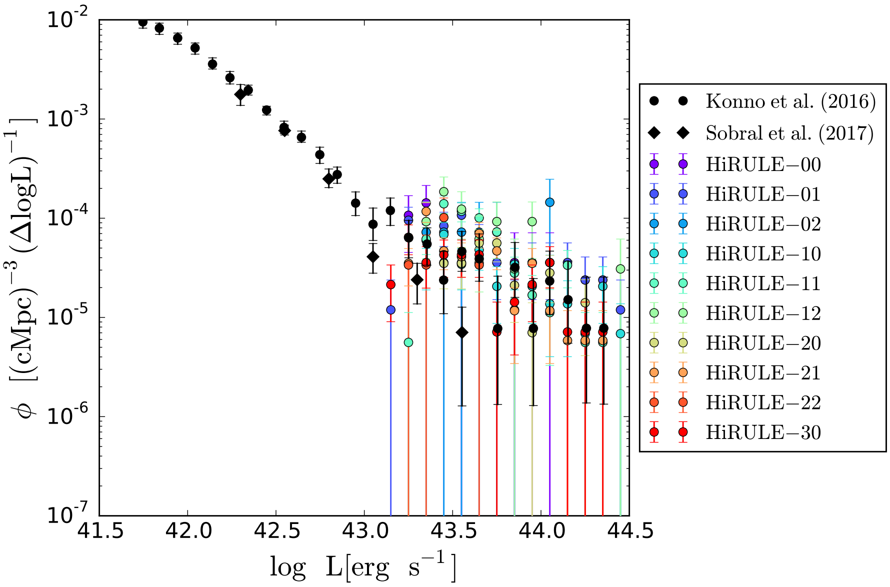


Fig. 6.5 Observed LAE luminosity function for each of the HiRULE field in different colors as labeled. For comparison we show the LF measured by Konno et al. (2016) and Sobral et al. (2017a) at the same redshift in black.

6.4 LAE Luminosity function

In this section we compute the LAE luminosity function in HiRULE and we compare to other surveys.

We compute the measured Lyα flux simply as

$$F_{\mathrm{Ly}\alpha} = \mathrm{FWHM}_{\mathrm{J0395}} \times (F_{\lambda,\mathrm{J0395}} - F_{\lambda,\mathrm{g}}), \tag{6.2}$$

where FWHM$_{\mathrm{J0395}}$ is the full width half maximum of the J0395 filter, $F_{\lambda,\mathrm{J0395}}$ is the flux in units of wavelength measured in the J0395 filter and $F_{\lambda,\mathrm{g}}$ is the flux in units of wavelength measured in the g band of HSC. Then, the Lyα luminosity $L_{\mathrm{Ly}\alpha}$ is computed assuming that the source is at redshift 2.24, which corresponds to the redshift necessary to shift Lyα to the pivot wavelength of J0395.

First we present the LAE LF field by field in Fig. 6.5. The volume of each field has been computed taking into account the FWHM of the J0395 filter and the sky coverage of HiRULE in each field, as listed in Table 6.1. Then the error bars are computed

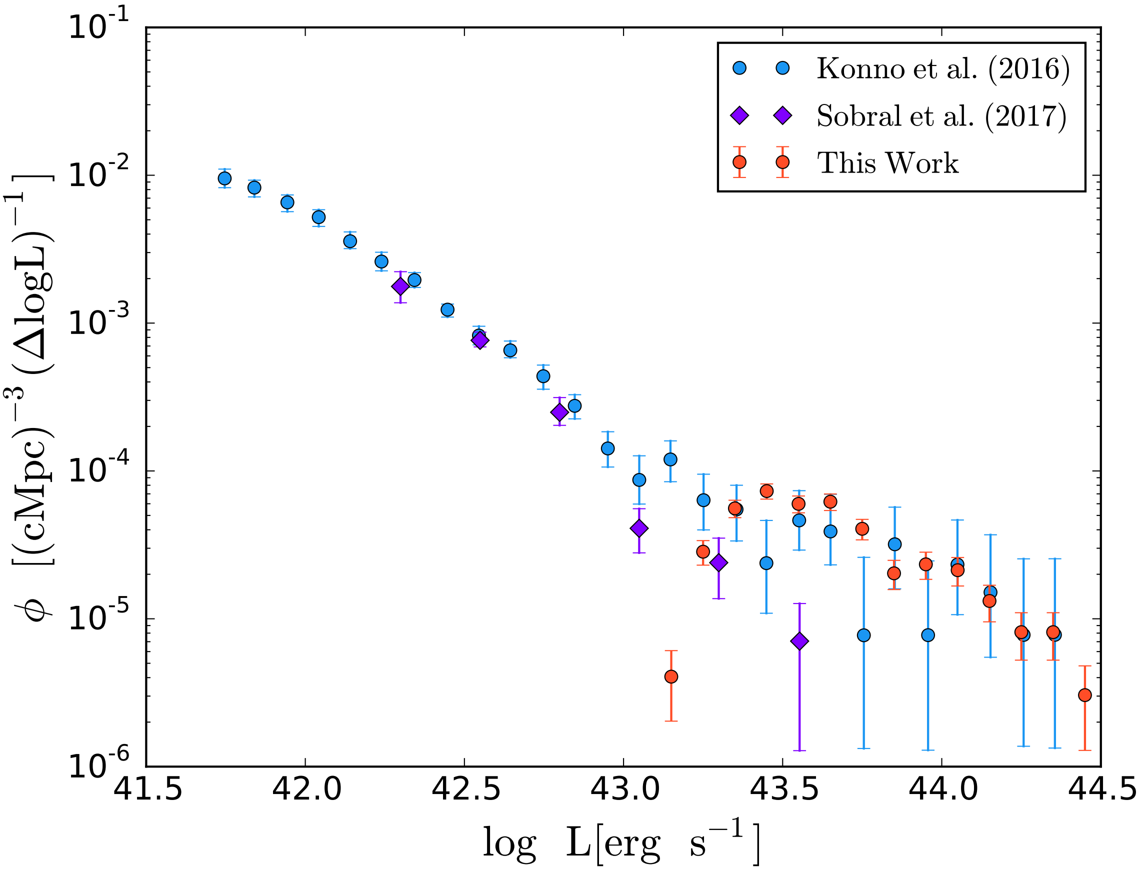


Fig. 6.6 Same as Fig 6.5 but for the whole LAE sample of LAEs in HiRULE.

assuming Poisson noise.

We compare our luminosity functions with the observations carried by Konno et al. (2016). HiRULE is pretty shallow in comparison with Konno et al. (2016) and Sobral et al. (2017a). HiRULE only reaches $L_{\mathrm{Ly}\alpha} \sim 10^{43.2} erg/s$. The size of the error bars on each of the individual field of HiRULE is comparable with the error bars of Konno et al. (2016) and Sobral et al. (2017a). Additionally, the HiRULE LF distributes as a cloud around observations.

the LAE luminosity function of HiRULE improves when all the field are combined. In Fig. 6.6 we shoe the LAE LF of the whole HiRULE volume. Thanks to the large volume scanned by HiRULE the error bars become smaller than the ones in Konno et al. (2016) and Sobral et al. (2017a). Additionally, our LF falls on top of the one measured by Konno et al. (2016). Meanwhile, our LF predicts a larger number of

LAEs at $\mathrm{L}_{\mathrm{Ly}\alpha} \sim 10^{43.5} erg/s$ than Sobral et al. (2017a). However, the error bars of Sobral et al. (2017a) LF at this luminosity are relatively big, so our measurements are not strongly incompatible. This suggested that our candidate selection is successful in retrieving LAEs. In this way, HiRULE confirms the already observed bright end of the LAE LF.

HiRULE's luminosity function drops very fast for $L_{\mathrm{Ly}\alpha} < 10^{43.5} erg/s$. This suggests that the completeness of our candidate selection starts decaying around the same Lyα luminosity.

6.5 Conclusions and future work

In this chapter we have described HiRULE (High Redshift Universe Lyman-α survEy), a program targeting Lyman-α emitters at redshift $z \sim 2.2$. The main goals of this project is to probe the interplay between QSO and LAEs. HiRULE is based on narrow band imaging. Up to now $20deg^2$ have been observed across 10 different fields.

The LAE candidate selection is based on the combination of the HSC (Aihara et al., 2018) broad bands and observations using the narrow band J0395 at T80Cam at the Observatorio Astrofísico de Javalambre. With this data set we perform mainly color cuts to retrieve a population of LAE candidates at $z \sim 2.2$.

We have computed a preliminary LAE luminosity function. But comparing our result with the measurements by Konno et al. (2016) and Sobral et al. (2017a) we reach the conclusion that our candidate selection works well up to $L_{\mathrm{Ly}\alpha} \sim 10^{43.5} erg/s$, where the number counts of LAE candidates in HiRULE drops. Indeed, our LAE LF agree very well with the result obtain by Konno et al. (2016) up to this luminosity.

HiRULE is at a very early stage of development. In the future, we plan to

1. Improve the measurements of $L_{\mathrm{Ly}\alpha}$ by using more sophisticated methods (e.g. Vilella-Rojo et al., 2015).
2. Extend the sky coverage of HiRULE.
3. Estimate the completeness and contamination of HiRULE.

Fig. 6.7 Logarithmic render of HiRULE-30 in J0395. The Field of view is $\sim 2\ deg^2$.

4. Compute the auto-correlation function of LAEs.
5. Compute the cross-correlation function between LAEs and QSO in the HiRULE footprint.
6. Study the Lyα florescence around the QSO found in HiRULE footprint.

Chapter 7

Summary, conclusions and future work

7.1 Summary and Conclusions

In this thesis we have tackled the problem of determining the role of properties of the interstellar and intergalactic medium in the Lyα transmission and how this reflects into the LAE observed population. We have explored various aspects of the radiative transfer that Lyα radiation experiences in astrophysical environments. Particular attention has been paid to the impact of these processes to the selection function of Lyα selected samples. The Lyα RT modulates the Lyα flux that reaches observers on Earth. In principle, the large cross-section between Lyα and atomic hydrogen causes Lyα to be very sensitive to any cloud of $\mathrm{H_I}$ encountered during its trajectory. This causes Lyα radiation to scatter inside both, the intergalactic and the interstellar medium. Therefore, the galaxies that we observed as Lyα emitters are sources for which Lyα photons manage to escape from the original galaxy and travel through the intergalactic medium unperturbed.

In order to study the selection function of LAEs due to the Lyα RT, we have developed a theoretical model that includes Lyα RT in the interstellar and intergalactic medium. Our model description begins at Chapter 2, where we explain how the Lyα RT in the interstellar medium can be implemented in big cosmological volumes. Our approach relies on the combination of a Radiative Transfer Monte Carlo code, the N-body simulation `P-Millennium` (Baugh et al., 2019) and the semi-analytical model of galaxy and formation `GALFORM` (Lacey et al., 2016). The `P-Millennium` simulation provides an accurate description of of the hierarchical growth of structures in the standard cosmological model (ΛCDM). `GALFORM` populates the simulation with galaxies following a fully-fledged, physical physically motivated set of differential equations that are able to provide a realistic description of different galaxy populations (Lacey et al., 2016).

Radiative Transfer Monte Carlo codes have proved to be an useful tool to understand the behaviour of Lyα radiation in the presence of $\mathrm{H_I}$ (e.g. Ahn et al., 2000; Zheng and Miralda-Escudé, 2002; Orsi et al., 2012). These algorithms simulate the radiative transfer processes that Lyα radiation suffer inside $\mathrm{H_I}$ clouds. In this way, properties such as the escape fraction of Lyα photons or the Lyα line profile can be studied for different gas configurations. These codes become very useful when studying the Lyα radiation that emerges from the ISM of galaxies. The gas geometry and properties of the simulated $\mathrm{H_I}$ clouds can be set to replicate the ISM structures. It is of particular importance the study of outflows. Because of their expansion velocity, radiation initially emitted at Lyα is received Doppler shifted, which facilitates its escape. We have built a big sample of simulation with `LyaRT`, the radiative transfer Monte Carlo code introduced by Orsi et al. (2012). From this grid of simulations we derived analytical expressions to compute the Lyα escape fraction for different outflow configurations in a every efficient way.

Afterwards, we link the outflow properties to the properties of each galaxy provided by `GALFORM`. In this way, we assign to every simulated galaxy a different ISM escape fraction that depends on its unique properties. This allows us to study the selection function of Lyα selected samples in a large cosmological volume. We find that LAEs constitute a very particular galaxy population. LAEs are galaxies with intermediate mass and, more importantly, they are metal-poor and exhibit moderate star formation. The intrinsic Lyα radiation generated in regions of star formation increases with the star formation rate. However, due to the SFR-metallicity relation, galaxies with the strongest SFR are also the most metal-rich. Additionally, Lyα radiation is strongly attenuated by dust due to its RT. This creates a balance between metallicity (Z) and SFR in Lyα selected samples. For the galaxies with the highest Z (and SFR) the Lyα radiation does not manage to escape the ISM, making theses galaxies not visible as LAEs. Metallicity and SFR are thus the main properties controling the Lyα radiation that escapes from the galaxy. However, we find that the mass of the galaxy also plays an important role. In particular, less massive galaxies have a higher escape fraction, but a lower intrinsic Lyα emission. This causes galaxies with moderate mass to be more likely observed as LAEs. This has a direct impact in the clustering amplitude of

LAEs.

The next step to build a comprehensive model of LAEs is the implementation of the Lyα RT in the intergalactic medium (IGM). In order to make this upgrade we need to include several new pieces to our model. As the IGM transmission depends on the frequency of the photons, it becomes crucial to model properly the line profile of LAEs when implementing the RT inside the IGM. For this we design `FLaREON` in Chapter 3, a public `Python` software that predicts a large number of Lyα escape fractions and line profiles in a short amount of time. Our new tool is based on the Monte Carlo radiative transfer code `LyaRT`. In order to save computational time, `FLaREON` takes a completely different approach and makes use of a set of pre-computed property grids for different outflow geometries with `LyaRT`. Then it uses machine learning and multidimensional interpolation techniques to predict the Lyα escape fraction and line profile for a arbitrary outflow configuration. This decreases more than 7 orders of magnitude the computational time while the accuracy remains excellent. Additionally, `FLaREON` is not only valid for populating large simulation, but it is a great tool to understand observations. In particular, it is very useful to constrain outflow properties from observed Lyα line profiles.

The other key piece that we need is the IGM transmission curves for every location of our simulation. We computed this from the density field, velocity field and the ionizing background of our simulation. The density and velocity field are calculated form the dark matter halos. The ionization rate is computed as the superposition of the ionizing radiation emitted by each galaxy. In this way, the UV background at a given position depends on its environment. In practice, IGM regions close to big sources of ionizing photons have a greater ionization rate.

All together, we associate an ISM Lyα escape fraction and line profile to each galaxy depending on its own properties. Then, we convolve the galaxy Lyα line profile with the IGM transmission along the line of sight for that particular galaxy. This gives the escape fraction of Lyα from the IGM. The whole model can thus be used to make mocks of realistic Lyα flux that an observer could detect.

Analyzing the effects of the IGM at different comic times, we first focus on the epoch of post-reionization, when the IGM is almost completely ionized in Chapter 4. We find that the selection function on galaxy properties is not affected by the RT in the

IGM in Lyα selected samples. However, IGM properties modulate the observability of Lyα, as the IGM transmission is enhanced in low density environments and in regions where is more probable galaxies to be distancing the IGM. This causes a coupling between the Lyα observability and the large scale IGM properties, as first studied by Zheng et al. (2010, 2011). We find that the RT inside the IGM causes selection effects on the IGM density ρ, its gradient $\partial_{\rm z}\rho$, the IGM velocity field along the of sight $\rm V_z$ and its gradient $\partial_{\rm z}\rm V_z$. The amplitude of the coupling depends on redshift. At high redshift the Universe is denser, which translates into a larger absorption of Lyα photons and a greater coupling. In general, the amplitude of the trends we find is small. In particular, the IGM transmission only varies a few percent at the highest redshift studied here.

The coupling between Lyα observability and the IGM large scale properties has a significant impact in the measured clustering of Lyα selected samples (Zheng et al., 2011). The coupling of the different IGM properties have different effects in the measured clustering of LAEs. Our model predicts that only the coupling with the IGM velocity along the line of sight is strong enough to notably modify the clustering of LAEs. At high redshift the clustering of LAEs is distorted. Indeed, the clustering on large scales is boosted non-linearly. This becomes crucial at large scales. In order to study the clustering of LAE on this scales we fit a clustering model that includes the coupling with the IGM (Zheng et al., 2011) to our clustering predictions on the scale range where they are reliable. Then, we extrapolate and find that the BAO peak shape is also distorted and its maximum is shifted by up to $\sim$ 1cMpc/h. Finally, we also study the clustering of LAEs along and perpendicular to the line of sight. We find that the IGM coupling is not strong enough to induce any asymmetry between these two clustering components. Overall, this highlights the importance of including Lyα RT in cosmological studies of LAEs.

In Chapter 5 we move our model to the epoch of reionization. The first main difference with respect to the previous implementation is the topology of the local fraction of neutral hydrogen during this epoch. We mimic the $\rm H_I$ distribution in the IGM by considering self-shielded to ionizing radiation (thus, neutral) regions below a certain photoionization rate threshold. This threshold and the global fraction of neutral $\langle \rm X_{H_I} \rangle$ hydrogen are directly linked. The higher the threshold is, the larger is the volume considered neutral and the larger is $\langle \rm X_{H_I} \rangle$.

The second feature of our model during the epoch of reionization is that the redshift depends on the line of sight distance. Thanks to this we can assign different global fraction of neutral hydrogen to a given distance along the line of sight. Indeed, we use a simple toy-model for generating histories of reionization, which depends on when and how fast it was. Then, we compute the photoionization rate field for each of these histories. Finally, we compute the IGM transmission and convolve it with the Lyα emission lines emerging from galaxies.

The IGM transmission is lower in $\mathrm{H_I}$ clumps than in $\mathrm{H_{II}}$ bubbles. This causes several of the statistical properties of LAEs, such us the luminosity function and clustering, to change for each reionization history. In particular, we find that the number of observed LAEs depends strongly on the global fraction of neutral hydrogen. When $\langle \mathrm{X_{H_I}} \rangle$ is large, fewer LAEs are detected. This translates into a modification of the LAE luminosity function for each reionization history. In general, we find that the lower the global fraction of neutral hydrogen at the center of our box ($\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$), the lower the number counts of LAEs and the smaller the LF amplitude. For small values of $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$ we find that if reionization is fast, then less LAEs are found. However, if $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$ is high and reionization is fast, a greater number of LAEs is found in comparison to the case in which reionization is slow. This is due to the greater amount of ionized volume when reionization is fast for high values of $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$.

Our model predicts that a precise analysis of the clustering of LAEs can provide constrains on the history of reionization of the Universe. In particular, we find that the time and duration of reionization modify in a different way the clustering of LAEs. While the clustering of LAEs on small scales is affected by the value of $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$, on large scales the speed of reionization plays a great role. The amplitude of the clustering increases with $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$. Meanwhile, the shape of the clustering is distorted by how fast reionization takes place. Indeed, if reionization is fast, the 2-point correlation function of LAEs flattens on large scales.

We also divide the clustering in parallel and perpendicular components. We find that increasing $\langle \mathrm{X_{H_I}}(z = 6.7) \rangle$ boosts the clustering along the line of sight. Meanwhile, the faster reionization takes place, the more the clustering in the directions perpendicular to the line of sight is enhanced. In this way, the study of the cuadrupole of LAEs is also very useful to constrain the history of reionization.

Finally, we have presented the first results of HiRULE (High Redshift Universe Lyα survEy) in Chapter 6. The main goal of this survey is to study the cross-correlation between LAEs and QSO at redshift 2.2. For this we make use of photometric narrow band observations to measure Lyα and broad band photometry to measure the continuum of our candidates. The narrow band data comes from observations using T80Cam at the Observatio Astrofísico de Javalambre (OAJ). For the continuum detection we use the Hyper Suprime Cam (HSC) broad bands (Aihara et al., 2018). Therefore, HiRULE footprint is the overlap between the OAJ narrow band observations and HSC. The effective sky coverage of HiRULE is $\sim 9.3\ deg^2$. The detection threshold (signal to noise ratio = 5) in this field is magnitude ~ 22 in the narrow band. After reducing the data we make our LAE candidate selection. This selection is based mostly on color cuts. We recover a total of 707 LAE candidates all over the HiRULE footprint. We compute the observed luminosity function of our LAE candidates and we find a very good agreement with other works in the literature. HiRULE is at an early stage and we will continue working on it.

All together, we have focused our efforts on understanding the effects of the Lyα radiative transfer on Lyα selected samples. We conclude that the Lyα RT inside the interstellar and intergalactic medium determines which galaxies are observed as LAEs. One one hand, the Lyα RT inside galaxies selects metal-poor galaxies with moderate star formation rate. On the other hand, the RT in the IGM selects galaxies in regions that are more transparent. This modifies the clustering of LAEs during and after the reionization of the Universe. In the post-reionization epoch, the coupling of the Lyα observability and velocity causes a non-linear boost in the clustering at large scales. The modeling of this feature will be of extreme importance in the up coming cosmological surveys using LAEs as tracers of the large scale structure of the Universe. Meanwhile, during the epoch of reionization, the local fraction of neutral hydrogen topology the modifies the spatial distribution of LAEs. Therefore, the clustering of LAEs is a promising tool to constrain the history of reionization of the Universe.

7.2 Future work

In this thesis we have studied how to implement the Lyα radiative processes in cosmological simulations. The methodology presented here can be applied to other cosmological simulation with a modest computational cost. This opens a wide variety of possibilities to future projects.

Chapters 5 and 6 present work in early stages of development, and thus with great potential for exploring new avenues of research. Here we list some ideas to improve and develop the work presented in these chapters as immediate future work. Then, we also list some other ideas the future.

1. **Improving our model of the epoch reionization**: The model, as presented in Chapter 5 can be easily upgraded. As discussed, some over-simplistic assumptions can be easily revised.

   The redshift of the Universe evolve along the line of sight of our box. However, we assume that the comoving density field, velocity field and galaxy population are constant in our simulation box. Some evolution in each of these properties from $z = 5.7$ to 7.9 (the edges of our box) is expected. In particular, the comoving density field is smoother at higher redshifts, as structures collapse hierarchically. This also affects the galaxy population. Less galaxies are formed at higher redshifts. Additionally, galaxy properties, such as metallicity or stellar age, also evolve with redshifts.

   To overcome this caveat we plan to compute the IGM properties to all the snapshots available of `P-Millennium` between redshift 5.7 and 7.9. Then we will also run `GALFORM` for each of these snapshots. Finally, we will build a new simulation box. In this new box the properties of the IGM and galaxies will depend on the distance along the line of sight through its redshift dependence. In this way the IGM density field will be smoother at the further edge of the box. At the same time, the galaxy population will be more evolved at the close edge of the box.

   Overall, we expect that the main trends shown here remain after improving our model. Our models, however, will be more robust and the galaxy properties will be better determined. This will allow us to study the properties of those galaxies that even if they are hosted in $\mathrm{H_I}$ clumps, they are still observed as LAEs. These galaxies are very interesting, as they constitute the most extreme LAEs at a very early epoch of the Universe.

2. **Cross-correlation between LAEs and** $\mathrm{H_I}$ $21cm$ **emission**: Once our model at the epoch of reionization is refined we can also study the cross-correlation with $21cm$ emission. This radiation is produced by atomic hydrogen. In particular, the emission comes from transitions between hiperfine structure energy levels of the hydrogen atom (see Pritchard and Loeb, 2012, for a review). Therefore, its brightness is proportional to the neutral hydrogen number density.

   During the epoch of reionization a significant fraction of the Universe is in a neutral estate. This produces that the $21cm$ radiation is very strong and heavily correlated with $\mathrm{H_I}$ clumps. Meanwhile, at the same epoch, LAEs avoid neutral regions and tend to inhabit $\mathrm{H_{II}}$ bubbles. Therefore, the computation of their cross-correlation can cast additional light about the history of reionization of the Universe. The predictions of our model might have a great impact in the forthcoming SKA (Square Kilometer Array, Koopmans et al. (2015)) project, which will measure the $21cm$ radiation from the cosmic dawn to reionization.

3. **HiRULE**: Our survey is at an early development phase and there is much and more to improve and do. We plan to i) improve the measurements of $L_{\mathrm{Ly}\alpha}$ by using more sophisticated methods (e.g. Vilella-Rojo et al., 2015); ii) to extend the sky coverage of HiRULE; iii) estimate the completeness and contamination of HiRULE; iv) compute the angular auto-correlation function of LAEs; v) compute the cross-correlation function between LAEs and QSO in the HiRULE footprint; vi) and to study the Lyα florescence around the QSO found in HiRULE footprint. We expect our measurements to complement those produced by the HETDEX survey (Hill et al., 2008a).

4. **Upgrade `FLaREON`**: Up to now most of the observational studies of Lyα line profiles have focused on the Lyα line profile integrating all the Lyα light of a source (e.g. Gronke, 2017). However, in the future, new instruments will allow us to study the Lyα line profile at different regions of the source. In this way it will be crucial to understand how the Lyα line profile evolves from inner to further regions.

   This can be easily studied with the radiative transfer Monte Carlo code `LyaRT`, the progenitor of `FLaREON`. Later on, we can use the same approach as in Chapter 3 to implement this into `FLaREON`. Some early results point out that the radial evolution of the line profile is quite complex and it depends on the gas geometry assumed. Therefore, a rigorous comparison between observational data and a proper model can help to unveil the outflow geometry driving the escape of Lyα photons.

5. **Cross-correlation between LAEs and QSO**: There are two main kinds of LAEs: i) star forming galaxies and ii) galaxies with an active nuclei (AGNs). Some observational studies have found that star forming LAEs populate the intermediate and low brightness range of the LAE LF, while AGNs selected by their Lyα populate the bright end (e.g. Konno et al., 2016).

   The Lyα production mechanism is different for each LAE population, as well as the interstellar environment. This causes differences between the properties of the Lyα that emerges from the galaxies and enters intro the IGM. In particular, the Lyα line profiles produced by outflows driven by star formation (as the those studied in this thesis) have up to a few angstroms of width (see Fig. 1.16). In comparison, the Lyα line profile of QSOs tends to be broader (about a few tens of angstroms) and more redshifted (Corbin, 1995).

   The differences in width and redshift of the Lyα line before entering into the IGM are crucial. As QSO Lyα lines are broader and more redshifted, in principle, they are less attenuated by the IGM. This would lower the coupling between IGM properties and Lyα observability. Basically, we expect that the AGN LAEs are almost independent to the IGM, even at the highest redshift, while star forming LAEs are coupled to their environment.

   Moreover, this would cause that LAEs with moderate Lyα luminosity and LAEs with bright Lyα flux inhabit slight different regions of the Universe, as star forming LAEs prefer regions with $\mathrm{V_z} > 0$. This would lower the cross-correlation amplitude between QSO and star forming LAEs (even if both are detected using Lyα). In particular, the greater is the coupling of the IGM with star forming LAEs, the lower should be the cross-correlation.

We plan to incorporate QSOs in our models to be able to make this study. Finally, if there is a lower coupling (or no coupling at all) between bright LAEs (QSOs) and the IGM large scale properties, then, the clustering of AGN LAEs would not be distorted at large scales. This would relax the complexity of extracting cosmological information from bright Lyα selected samples.

7.3 Final remarks

This thesis focus on the selection function of the galaxy population known as *Lyman-α emitters* at several cosmic epochs. Our predictions highlight the importance of including the Lyα radiative transfer in the analysis of cosmological surveys tracing the Universe with Lyman-α emitters. The methodology presented here includes the complex multi-scale problem of the Lyα radiative transfer by implementing, at the same time, scales of the interstellar and intergalactic medium. This approach will be very useful to produce realistic mocks for forthcoming surveys such as J-PAS and HETDEX. Moreover, as a product of our research we have designed `FLaREON`, an open source code that predicts the properties of the Lyα emerging from galactic outflows. `FLaREON` will assist researchers to interpret Lyα line observations. Additionally, we have demonstrated that the study of Lyman-α emitters during the epoch of reionization cloud expand our knowledge about the Universe in the future.

References

.

Ahn, S.
2003. Singly-Peaked P-Cygni Type Lyα From Starburst Galaxies. *Journal of Korean Astronomical Society*, 36:145–148.

Ahn, S.
2004. Singly Peaked Asymmetric Lyα from Starburst Galaxies. , 601:L25–L28.

Ahn, S.-H., H.-W. Lee, and H. M. Lee
2000. Ly Alpha Transfer in a Thick, Dusty, and Static Medium. *Journal of Korean Astronomical Society*, 33:29–36.

Aihara, H., N. Arimoto, R. Armstrong, S. Arnouts, N. A. Bahcall, S. Bickerton, J. Bosch, K. Bundy, P. L. Capak, J. H. H. Chan, M. Chiba, J. Coupon, E. Egami, M. Enoki, F. Finet, H. Fujimori, S. Fujimoto, H. Furusawa, J. Furusawa, T. Goto, A. Goulding, J. P. Greco, J. E. Greene, J. E. Gunn, T. Hamana, Y. Harikane, Y. Hashimoto, T. Hattori, M. Hayashi, Y. Hayashi, K. G. Hełminiak, R. Higuchi, C. Hikage, P. T. P. Ho, B.-C. Hsieh, K. Huang, S. Huang, H. Ikeda, M. Imanishi, A. K. Inoue, K. Iwasawa, I. Iwata, A. T. Jaelani, H.-Y. Jian, Y. Kamata, H. Karoji, N. Kashikawa, N. Katayama, S. Kawanomoto, I. Kayo, J. Koda, M. Koike, T. Kojima, Y. Komiyama, A. Konno, S. Koshida, Y. Koyama, H. Kusakabe, A. Leauthaud, C.-H. Lee, L. Lin, Y.-T. Lin, R. H. Lupton, R. Mandelbaum, Y. Matsuoka, E. Medezinski, S. Mineo, S. Miyama, H. Miyatake, S. Miyazaki, R. Momose, A. More, S. More, Y. Moritani, T. J. Moriya, T. Morokuma, S. Mukae, R. Murata, H. Murayama, T. Nagao, F. Nakata, M. Niida, H. Niikura, A. J. Nishizawa, Y. Obuchi, M. Oguri, Y. Oishi, N. Okabe, S. Okamoto, Y. Okura, Y. Ono, M. Onodera, M. Onoue, K. Osato, M. Ouchi, P. A. Price, T.-S. Pyo, M. Sako, M. Sawicki, T. Shibuya, K. Shimasaku, A. Shimono, M. Shirasaki, J. D. Silverman, M. Simet, J. Speagle, D. N. Spergel, M. A. Strauss, Y. Sugahara, N. Sugiyama, Y. Suto, S. H. Suyu, N. Suzuki, P. J. Tait, M. Takada, T. Takata, N. Tamura, M. M. Tanaka, M. Tanaka, M. Tanaka, Y. Tanaka, T. Terai, Y. Terashima, Y. Toba, N. Tominaga, J. Toshikawa, E. L. Turner, T. Uchida, H. Uchiyama, K. Umetsu, F. Uraguchi, Y. Urata, T. Usuda, Y. Utsumi, S.-Y. Wang, W.-H. Wang, K. C. Wong, K. Yabe, Y. Yamada, H. Yamanoi, N. Yasuda, S. Yeh, A. Yonehara, and S. Yuma
2018. The Hyper Suprime-Cam SSP Survey: Overview and survey design. , 70:S4.

Alam, S., M. Ata, S. Bailey, F. Beutler, D. Bizyaev, J. A. Blazek, A. S. Bolton, J. R. Brownstein, A. Burden, C.-H. Chuang, J. Comparat, A. J. Cuesta, K. S. Dawson, D. J.

Eisenstein, S. Escoffier, H. Gil-Marín, J. N. Grieb, N. Hand, S. Ho, K. Kinemuchi, D. Kirkby, F. Kitaura, E. Malanushenko, V. Malanushenko, C. Maraston, C. K. McBride, R. C. Nichol, M. D. Olmstead, D. Oravetz, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, M. Pellejero-Ibanez, W. J. Percival, P. Petitjean, F. Prada, A. M. Price-Whelan, B. A. Reid, S. A. Rodríguez-Torres, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martín, S. Saito, S. Salazar-Albornoz, L. Samushia, A. G. Sánchez, S. Satpathy, D. J. Schlegel, D. P. Schneider, C. G. Scóccola, H.-J. Seo, E. S. Sheldon, A. Simmons, A. Slosar, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. V. Magaña, J. A. Vazquez, L. Verde, D. A. Wake, Y. Wang, D. H. Weinberg, M. White, W. M. Wood-Vasey, C. Yèche, I. Zehavi, Z. Zhai, and G.-B. Zhao
2017. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. , 470:2617–2652.

Amendola, L. and S. Tsujikawa
2010. *Dark Energy: Theory and Observations.*

Angulo, R. E., V. Springel, S. D. M. White, A. Jenkins, C. M. Baugh, and C. S. Frenk
2012. Scaling relations for galaxy clusters in the Millennium-XXL simulation. , 426:2046–2062.

Barkana, R. and A. Loeb
2001. In the beginning: the first sources of light and the reionization of the universe. , 349:125–238.

Barnes, L. A. and M. G. Haehnelt
2010. Faint extended Lyα emission due to star formation at the centre of high column density QSO absorption systems. , 403:870–885.

Baugh, C. M.
2006. A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports on Progress in Physics*, 69:3101–3156.

Baugh, C. M., V. Gonzalez-Perez, C. del P Lagos, C. G. Lacey, J. C. Helly, A. Jenkins, C. S. Frenk, A. J. Benson, R. G. Bower, and S. Cole
2019. Galaxy formation in the Planck Millennium: the atomic hydrogen content of dark matter halos. .

Baugh, C. M., C. G. Lacey, C. S. Frenk, G. L. Granato, L. Silva, A. Bressan, A. J. Benson, and S. Cole
2005. Can the faint submillimetre galaxies be explained in the Λ cold dark matter model? , 356:1191–1200.

Behrens, C., C. Byrohl, S. Saito, and J. C. Niemeyer
2017. The impact of Lyman-α radiative transfer on large-scale clustering in the Illustris simulation. *ArXiv e-prints.*

Behrens, C., A. Pallottini, A. Ferrara, S. Gallerani, and L. Vallini
2019. Ly α emission from galaxies in the Epoch of Reionization. , 486:2197–2209.

Behroozi, P. S., C. Conroy, and R. H. Wechsler
2010. A Comprehensive Analysis of Uncertainties Affecting the Stellar Mass-Halo Mass Relation for 0 z 4. , 717:379–403.

Benitez, N., R. Dupke, M. Moles, L. Sodre, J. Cenarro, A. Marin-Franch, K. Taylor, D. Cristobal, A. Fernandez-Soto, C. Mendes de Oliveira, J. Cepa-Nogue, L. R. Abramo, J. S. Alcaniz, R. Overzier, C. Hernandez-Monteagudo, E. J. Alfaro, A. Kanaan, J. M. Carvano, R. R. R. Reis, E. Martinez Gonzalez, B. Ascaso, F. Ballesteros, H. S. Xavier, J. Varela, A. Ederoclite, H. Vazquez Ramio, T. Broadhurst, E. Cypriano, R. Angulo, J. M. Diego, A. Zandivarez, E. Diaz, P. Melchior, K. Umetsu, P. F. Spinelli, A. Zitrin, D. Coe, G. Yepes, P. Vielva, V. Sahni, A. Marcos-Caballero, F. Shu Kitaura, A. L. Maroto, M. Masip, S. Tsujikawa, S. Carneiro, J. Gonzalez Nuevo, G. C. Carvalho, M. J. Reboucas, J. C. Carvalho, E. Abdalla, A. Bernui, C. Pigozzo, E. G. M. Ferreira, N. Chandrachani Devi, C. A. P. Bengaly, Jr., M. Campista, A. Amorim, N. V. Asari, A. Bongiovanni, S. Bonoli, G. Bruzual, N. Cardiel, A. Cava, R. Cid Fernandes, P. Coelho, A. Cortesi, R. G. Delgado, L. Diaz Garcia, J. M. R. Espinosa, E. Galliano, J. I. Gonzalez-Serrano, J. Falcon-Barroso, J. Fritz, C. Fernandes, J. Gorgas, C. Hoyos, Y. Jimenez-Teja, J. A. Lopez-Aguerri, C. Lopez-San Juan, A. Mateus, A. Molino, P. Novais, A. OMill, I. Oteo, P. G. Perez-Gonzalez, B. Poggianti, R. Proctor, E. Ricciardelli, P. Sanchez-Blazquez, T. Storchi-Bergmann, E. Telles, W. Schoennell, N. Trujillo, A. Vazdekis, K. Viironen, S. Daflon, T. Aparicio-Villegas, D. Rocha, T. Ribeiro, M. Borges, S. L. Martins, W. Marcolino, D. Martinez-Delgado, M. A. Perez-Torres, B. B. Siffert, M. O. Calvao, M. Sako, R. Kessler, A. Alvarez-Candal, M. De Pra, F. Roig, D. Lazzaro, J. Gorosabel, R. Lopes de Oliveira, G. B. Lima-Neto, J. Irwin, J. F. Liu, E. Alvarez, I. Balmes, S. Chueca, M. V. Costa-Duarte, A. A. da Costa, M. L. L. Dantas, A. Y. Diaz, J. Fabregat, F. Ferrari, B. Gavela, S. G. Gracia, N. Gruel, J. L. L. Gutierrez, R. Guzman, J. D. Hernandez-Fernandez, D. Herranz, L. Hurtado-Gil, F. Jablonsky, R. Laporte, L. L. Le Tiran, J. Licandro, M. Lima, E. Martin, V. Martinez, J. J. C. Montero, P. Penteado, C. B. Pereira, V. Peris, V. Quilis, M. Sanchez-Portal, A. C. Soja, E. Solano, J. Torra, and L. Valdivielso
2014a. J-PAS: The Javalambre-Physics of the Accelerated Universe Astrophysical Survey. *ArXiv e-prints*.

Benitez, N., R. Dupke, M. Moles, L. Sodre, J. Cenarro, A. Marin-Franch, K. Taylor, D. Cristobal, A. Fernandez-Soto, C. Mendes de Oliveira, J. Cepa-Nogue, L. R. Abramo, J. S. Alcaniz, R. Overzier, C. Hernandez-Monteagudo, E. J. Alfaro, A. Kanaan, J. M. Carvano, R. R. R. Reis, E. Martinez Gonzalez, B. Ascaso, F. Ballesteros, H. S. Xavier, J. Varela, A. Ederoclite, H. Vazquez Ramio, T. Broadhurst, E. Cypriano, R. Angulo, J. M. Diego, A. Zandivarez, E. Diaz, P. Melchior, K. Umetsu, P. F. Spinelli, A. Zitrin, D. Coe, G. Yepes, P. Vielva, V. Sahni, A. Marcos-Caballero, F. Shu Kitaura, A. L. Maroto, M. Masip, S. Tsujikawa, S. Carneiro, J. Gonzalez Nuevo, G. C. Carvalho, M. J. Reboucas, J. C. Carvalho, E. Abdalla, A. Bernui, C. Pigozzo, E. G. M. Ferreira, N. Chandrachani Devi, C. A. P. Bengaly, Jr., M. Campista, A. Amorim, N. V. Asari, A. Bongiovanni, S. Bonoli, G. Bruzual, N. Cardiel, A. Cava, R. Cid Fernandes, P. Coelho, A. Cortesi, R. G. Delgado, L. Diaz Garcia, J. M. R. Espinosa, E. Galliano, J. I. Gonzalez-Serrano, J. Falcon-Barroso, J. Fritz, C. Fernandes, J. Gorgas, C. Hoyos, Y. Jimenez-Teja, J. A. Lopez-Aguerri, C. Lopez-San Juan, A. Mateus, A. Molino, P. Novais, A. OMill, I. Oteo, P. G.

Perez-Gonzalez, B. Poggianti, R. Proctor, E. Ricciardelli, P. Sanchez-Blazquez, T. Storchi-Bergmann, E. Telles, W. Schoennell, N. Trujillo, A. Vazdekis, K. Viironen, S. Daflon, T. Aparicio-Villegas, D. Rocha, T. Ribeiro, M. Borges, S. L. Martins, W. Marcolino, D. Martinez-Delgado, M. A. Perez-Torres, B. B. Siffert, M. O. Calvao, M. Sako, R. Kessler, A. Alvarez-Candal, M. De Pra, F. Roig, D. Lazzaro, J. Gorosabel, R. Lopes de Oliveira, G. B. Lima-Neto, J. Irwin, J. F. Liu, E. Alvarez, I. Balmes, S. Chueca, M. V. Costa-Duarte, A. A. da Costa, M. L. L. Dantas, A. Y. Diaz, J. Fabregat, F. Ferrari, B. Gavela, S. G. Gracia, N. Gruel, J. L. L. Gutierrez, R. Guzman, J. D. Hernandez-Fernandez, D. Herranz, L. Hurtado-Gil, F. Jablonsky, R. Laporte, L. L. Le Tiran, J. Licandro, M. Lima, E. Martin, V. Martinez, J. J. C. Montero, P. Penteado, C. B. Pereira, V. Peris, V. Quilis, M. Sanchez-Portal, A. C. Soja, E. Solano, J. Torra, and L. Valdivielso
2014b. J-PAS: The Javalambre-Physics of the Accelerated Universe Astrophysical Survey. *ArXiv e-prints.*

Bennett, C. L., D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright
2013. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. , 208:20.

Bielby, R. M., P. Tummuangpak, T. Shanks, H. Francke, N. H. M. Crighton, E. Bañados, J. González-López, L. Infante, and A. Orsi
2016. The VLT LBG redshift survey - V. Characterizing the z = 3.1 Lyman α emitter population. , 456:4061–4080.

Bower, R. G., A. J. Benson, R. Malbon, J. C. Helly, C. S. Frenk, C. M. Baugh, S. Cole, and C. G. Lacey
2006. Breaking the hierarchy of galaxy formation. , 370:645–655.

Brasken, M. and E. Kyrola
1998. Resonance scattering of Lyman alpha from interstellar hydrogen. , 332:732–738.

Burgess, A.
1965. Tables of hydrogenic photoionization cross-sections and recombination coefficients. , 69:1.

Cantalupo, S., C. Porciani, and S. J. Lilly
2008. Mapping Neutral Hydrogen during Reionization with the Lyα Emission from Quasar Ionization Fronts. , 672:48–58.

Caruana, J., L. Wisotzki, E. C. Herenz, J. Kerutt, T. Urrutia, K. B. Schmidt, R. Bouwens, J. Brinchmann, S. Cantalupo, M. Carollo, C. Diener, A. Drake, T. Garel, R. A. Marino, J. Richard, R. Saust, J. Schaye, and A. Verhamme
2018. The MUSE-Wide survey: a measurement of the Ly α emitting fraction among z 3 galaxies. , 473:30–37.

Cassata, P., O. Le Fèvre, B. Garilli, D. Maccagni, V. Le Brun, M. Scodeggio, L. Tresse, O. Ilbert, G. Zamorani, O. Cucciati, T. Contini, R. Bielby, Y. Mellier, H. J. McCracken, A. Pollo, A. Zanichelli, S. Bardelli, A. Cappi, L. Pozzetti, D. Vergani, and

E. Zucca
2011. The VIMOS VLT Deep Survey: star formation rate density of Lyα emitters from a sample of 217 galaxies with spectroscopic redshifts $2 \leq z \leq 6.6., 525 : A143$.

Cazzoli, S., S. Arribas, R. Maiolino, and L. Colina
2016. Neutral gas outflows in nearby [U]LIRGs via optical NaD feature. , 590:A125.

Cenarro, A. J., M. Moles, D. Cristóbal-Hornillos, A. Marín-Franch, A. Ederoclite, J. Varela, C. López-Sanjuan, C. Hernández-Monteagudo, R. E. Angulo, H. Vázquez Ramió, K. Viironen, S. Bonoli, A. A. Orsi, G. Hurier, I. San Roman, N. Greisel, G. Vilella-Rojo, L. A. Díaz-García, R. Logroño-García, S. Gurung-López, D. Spinoso, D. Izquierdo-Villalba, J. A. L. Aguerri, C. Allende Prieto, C. Bonatto, J. M. Carvano, A. L. Chies-Santos, S. Daflon, R. A. Dupke, J. Falcón-Barroso, D. R. Gonçalves, Y. Jiménez-Teja, A. Molino, V. M. Placco, E. Solano, D. D. Whitten, J. Abril, J. L. Antón, R. Bello, S. Bielsa de Toledo, J. Castillo-Ramírez, S. Chueca, T. Civera, M. C. Díaz-Martín, M. Domínguez-Martínez, J. Garzarán-Calderaro, J. Hernández-Fuertes, R. Iglesias-Marzoa, C. Iñiguez, J. M. Jiménez Ruiz, K. Kruuse, J. L. Lamadrid, N. Lasso-Cabrera, G. López-Alegre, A. López-Sainz, N. Maícas, A. Moreno-Signes, D. J. Muniesa, S. Rodríguez-Llano, F. Rueda-Teruel, S. Rueda-Teruel, I. Soriano-Laguía, V. Tilve, L. Valdivielso, A. Yanes-Díaz, J. S. Alcaniz, C. Mendes de Oliveira, L. Sodré, P. Coelho, R. Lopes de Oliveira, A. Tamm, H. S. Xavier, L. R. Abramo, S. Akras, E. J. Alfaro, A. Alvarez-Candal, B. Ascaso, M. A. Beasley, T. C. Beers, M. Borges Fernandes, G. R. Bruzual, M. L. Buzzo, J. M. Carrasco, J. Cepa, A. Cortesi, M. V. Costa-Duarte, M. De Prá, G. Favole, A. Galarza, L. Galbany, K. Garcia, R. M. González Delgado, J. I. González-Serrano, L. A. Gutiérrez-Soto, J. A. Hernandez-Jimenez, A. Kanaan, H. Kuncarayakti, R. C. G. Landim, J. Laur, J. Licandro, G. B. Lima Neto, J. D. Lyman, J. Maíz Apellániz, J. Miralda-Escudé, D. Morate, J. P. Nogueira-Cavalcante, P. M. Novais, M. Oncins, I. Oteo, R. A. Overzier, C. B. Pereira, A. Rebassa-Mansergas, R. R. R. Reis, F. Roig, M. Sako, N. Salvador-Rusiñol, L. Sampedro, P. Sánchez-Blázquez, W. A. Santos, L. Schmidtobreick, B. B. Siffert, E. Telles, and J. M. Vilchez
2019. J-PLUS: The Javalambre Photometric Local Universe Survey. , 622:A176.

Chaves-Montero, J., R. E. Angulo, and C. Hernández-Monteagudo
2018. The effect of photometric redshift uncertainties on galaxy clustering and baryonic acoustic oscillations. , 477:3892–3909.

Chisholm, J., I. Orlitová, D. Schaerer, A. Verhamme, G. Worseck, Y. I. Izotov, T. X. Thuan, and N. G. Guseva
2017. Do galaxies that leak ionizing photons have extreme outflows? , 605:A67.

Cole, S., C. G. Lacey, C. M. Baugh, and C. S. Frenk
2000. Hierarchical galaxy formation. , 319:168–204.

Contreras, S., I. Zehavi, N. Padilla, C. M. Baugh, E. Jiménez, and I. Lacerna
2019. The evolution of assembly bias. , 484:1133–1148.

Cooray, A., J. Bock, D. Burgarella, R. Chary, T.-C. Chang, O. Doré, G. Fazio, A. Ferrara, Y. Gong, M. Santos, M. Silva, and M. Zemcov
2016. Cosmic Dawn Intensity Mapper. *arXiv e-prints*.

Corbin, M. R.
1995. QSO Broad Emission Line Asymmetries: Evidence of Gravitational Redshift? , 447:496.

Cowie, L. L. and E. M. Hu
1998. High-z Lyalpha Emitters. I. A Blank-Field Search for Objects near Redshift Z = 3.4 in and around the Hubble Deep Field and the Hawaii Deep Field SSA 22. , 115:1319–1328.

Dayal, P., A. Maselli, and A. Ferrara
2011. The visibility of Lyman α emitters during reionization. , 410:830–843.

DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen, C. Allende Prieto, J. Annis, S. Bailey, C. Balland, and et al.
2016. The DESI Experiment Part I: Science,Targeting, and Survey Design. *arXiv e-prints*.

Dijkstra, M.
2014. LyÎ Emitting Galaxies as a Probe of Reionisation. *Publications of the Astronomical Society of Australia*, 31:40.

Dijkstra, M.
2017. Saas-Fee Lecture Notes: Physics of Lyman Alpha Radiative Transfer. *ArXiv e-prints*.

Dijkstra, M., Z. Haiman, and M. Spaans
2006. Lyα Radiation from Collapsing Protogalaxies. I. Characteristics of the Emergent Spectrum. , 649:14–36.

Dijkstra, M., A. Lidz, and J. S. B. Wyithe
2007. The impact of The IGM on high-redshift Lyα emission lines. , 377:1175–1186.

Dijkstra, M. and A. Loeb
2009. Lyα blobs as an observational signature of cold accretion streams into galaxies. , 400:1109–1120.

Dijkstra, M., A. Mesinger, and J. S. B. Wyithe
2011. The detectability of Lyα emission from galaxies during the epoch of reionization. , 414:2139–2147.

Foreman-Mackey, D., D. W. Hogg, D. Lang, and J. Goodman
2013a. emcee: The MCMC Hammer. , 125:306.

Foreman-Mackey, D., D. W. Hogg, D. Lang, and J. Goodman
2013b. emcee: The MCMC Hammer. , 125:306.

Garcia-Vergara, C., J. F. Hennawi, L. F. Barrientos, and F. Arrigoni Battaia
2019. Clustering of Lyman-alpha Emitters Around Quasars at $z \sim 4$. *arXiv e-prints*.

García-Vergara, C., J. F. Hennawi, L. F. Barrientos, and H.-W. Rix
2017. Strong Clustering of Lyman Break Galaxies around Luminous Quasars at Z 4. , 848:7.

Garel, T., J. Blaizot, B. Guiderdoni, L. Michel-Dansac, M. Hayes, and A. Verhamme
2015. The UV, Lyman α, and dark matter halo properties of high-redshift galaxies. , 450:1279–1294.

Garel, T., J. Blaizot, B. Guiderdoni, D. Schaerer, A. Verhamme, and M. Hayes
2012. Modelling high redshift Lyman α emitters. , 422:310–325.

Gawiser, E., H. Francke, K. Lai, K. Schawinski, C. Gronwall, R. Ciardullo, R. Quadri, A. Orsi, L. F. Barrientos, G. A. Blanc, G. Fazio, J. J. Feldmeier, J. Huang, L. Infante, P. Lira, N. Padilla, E. N. Taylor, E. Treister, C. M. Urry, P. G. van Dokkum, and S. N. Virani
2007. Lyα-Emitting Galaxies at z = 3.1: L* Progenitors Experiencing Rapid Star Formation. , 671:278–284.

Gawiser, E., P. G. van Dokkum, D. Herrera, J. Maza, F. J. Castander, L. Infante, P. Lira, R. Quadri, R. Toner, E. Treister, C. M. Urry, M. Altmann, R. Assef, D. Christlein, P. S. Coppi, M. F. Durán, M. Franx, G. Galaz, L. Huerta, C. Liu, S. López, R. Méndez, D. C. Moore, M. Rubio, M. T. Ruiz, S. Toft, and S. K. Yi
2006. The Multiwavelength Survey by Yale-Chile (MUSYC): Survey Design and Deep Public UBVRIz' Images and Catalogs of the Extended Hubble Deep Field-South. , 162:1–19.

Geach, J. E., I. Smail, P. N. Best, J. Kurk, M. Casali, R. J. Ivison, and K. Coppin
2008. HiZELS: a high-redshift survey of Hα emitters - I. The cosmic star formation rate and clustering at z = 2.23. , 388:1473–1486.

Granato, G. L., C. G. Lacey, L. Silva, A. Bressan, C. M. Baugh, S. Cole, and C. S. Frenk
2000. The Infrared Side of Galaxy Formation. I. The Local Universe in the Semianalytical Framework. , 542:710–730.

Griffiths, D. J.
1982. *Introduction to quantum mechanics.*

Gronke, M.
2017. Modeling 237 Lyman-α spectra of the MUSE-Wide survey. , 608:A139.

Gronke, M., M. Dijkstra, M. McCourt, and S. P. Oh
2016. From Mirrors to Windows: Lyman-alpha Radiative Transfer in a Very Clumpy Medium. , 833:L26.

Gronwall, C., R. Ciardullo, T. Hickey, E. Gawiser, J. J. Feldmeier, P. G. van Dokkum, C. M. Urry, D. Herrera, B. D. Lehmer, L. Infante, A. Orsi, D. Marchesini, G. A. Blanc, H. Francke, P. Lira, and E. Treister
2007. Lyα Emission-Line Galaxies at z = 3.1 in the Extended Chandra Deep Field-South. , 667:79–91.

Guaita, L., E. Gawiser, N. Padilla, H. Francke, N. A. Bond, C. Gronwall, R. Ciardullo, J. J. Feldmeier, S. Sinawa, G. A. Blanc, and S. Virani
2010. Lyα-emitting Galaxies at z = 2.1 in ECDF-S: Building Blocks of Typical Present-day Galaxies? , 714:255–269.

Guaita, L., M. Talia, L. Pentericci, A. Verhamme, P. Cassata, B. C. Lemaux, I. Orlitova, B. Ribeiro, D. Schaerer, G. Zamorani, B. Garilli, V. Le Brun, O. Le Fèvre, D. Maccagni, L. A. M. Tasca, R. Thomas, E. Vanzella, E. Zucca, R. Amorin, S. Bardelli, M. Castellano, A. Grazian, N. P. Hathi, A. Koekemoer, and F. Marchi
2017. The VIMOS Ultra Deep Survey. The role of HI kinematics and HI column density on the escape of Lyα photons in star-forming galaxies at 2 z 4. , 606:A19.

Gurung-Lopez, S., A. A. Orsi, and S. Bonoli
2018. FLaREON : a fast computation of Lyα escape fractions and line profiles. *ArXiv e-prints*.

Gurung López, S., Á. A. Orsi, S. Bonoli, C. M. Baugh, and C. G. Lacey
2018. Ly*alpha* emitters in a cosmological volume I: the impact of radiative transfer. *ArXiv e-prints*.

Guth, A. H.
1981. Inflationary universe: A possible solution to the horizon and flatness problems. , 23:347–356.

Haardt, F. and P. Madau
2012. Radiative Transfer in a Clumpy Universe. IV. New Synthesis Models of the Cosmic UV/X-Ray Background. , 746:125.

Hagen, A., R. Ciardullo, C. Gronwall, V. Acquaviva, J. Bridge, G. R. Zeimann, G. A. Blanc, N. A. Bond, S. L. Finkelstein, M. Song, E. Gawiser, D. B. Fox, H. Gebhardt, A. I. Malz, D. P. Schneider, N. Drory, K. Gebhardt, and G. J. Hill
2014. Spectral Energy Distribution Fitting of HETDEX Pilot Survey Lyα Emitters in COSMOS and GOODS-N. , 786:59.

Harrington, J. P.
1973. The scattering of resonance-line radiation in the limit of large optical depth. , 162:43–+.

Henry, A., D. A. Berg, C. Scarlata, A. Verhamme, and D. Erb
2018. A Close Relationship between Lyα and Mg II in Green Pea Galaxies. , 855:96.

Henyey, L. G. and J. L. Greenstein
1941. Diffuse radiation in the Galaxy. , 93:70–83.

Herenz, E. C., T. Urrutia, L. Wisotzki, J. Kerutt, R. Saust, M. Werhahn, K. B. Schmidt, J. Caruana, C. Diener, R. Bacon, J. Brinchmann, J. Schaye, M. Maseda, and P. M. Weilbacher
2017. The MUSE-Wide survey: A first catalogue of 831 emission line galaxies. , 606:A12.

Hill, G. J., K. Gebhardt, E. Komatsu, N. Drory, P. J. MacQueen, J. Adams, G. A. Blanc, R. Koehler, M. Rafal, M. M. Roth, A. Kelz, C. Gronwall, R. Ciardullo, and D. P. Schneider
2008a. The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results. In *Astronomical Society of the Pacific Conference Series*, T. Kodama, T. Yamada, & K. Aoki, ed., volume 399 of *Astronomical Society of the Pacific Conference Series*, Pp. 115–+.

Hill, G. J., K. Gebhardt, E. Komatsu, N. Drory, P. J. MacQueen, J. Adams, G. A. Blanc, R. Koehler, M. Rafal, M. M. Roth, A. Kelz, C. Gronwall, R. Ciardullo, and D. P. Schneider
2008b. The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results. In *Panoramic Views of Galaxy Formation and Evolution*, T. Kodama, T. Yamada, and K. Aoki, eds., volume 399 of *Astronomical Society of the Pacific Conference Series*, P. 115.

Hjerting, F.
1938. Tables Facilitating the Calculation of Line Absorption Coefficients. , 88:508–+.

Hu, E. M., L. L. Cowie, and R. G. McMahon
1998. The Density of Ly alpha Emitters at Very High Redshift. , 502:L99+.

Hubble, E.
1929. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173.

Inoue, A. K., K. Hasegawa, T. Ishiyama, H. Yajima, I. Shimizu, M. Umemura, A. Konno, Y. Harikane, T. Shibuya, M. Ouchi, K. Shimasaku, Y. Ono, H. Kusakabe, R. Higuchi, and C.-H. Lee
2018. SILVERRUSH. VI. A simulation of Lyα emitters in the reionization epoch and a comparison with Subaru Hyper Suprime-Cam survey early data. *ArXiv e-prints*.

Jeeson-Daniel, A., B. Ciardi, U. Maio, M. Pierleoni, M. Dijkstra, and A. Maselli
2012. Effect of intergalactic medium on the observability of Lyα emitters during cosmic reionization. , 424(3):2193–2212.

Jose, C., R. Srianand, and K. Subramanian
2013. Clustering at high redshift: the connection between Lyman α emitters and Lyman break galaxies. , 435:368–377.

Kaiser, N.
1987. Clustering in real space and in redshift space. , 227:1–21.

Kashikawa, N., K. Shimasaku, M. A. Malkan, M. Doi, Y. Matsuda, M. Ouchi, Y. Taniguchi, C. Ly, T. Nagao, M. Iye, K. Motohara, T. Murayama, K. Murozono, K. Nariai, K. Ohta, S. Okamura, T. Sasaki, Y. Shioya, and M. Umemura
2006. The End of the Reionization Epoch Probed by Lyα Emitters at z = 6.5 in the Subaru Deep Field. , 648:7–22.

Kashikawa, N., K. Shimasaku, Y. Matsuda, E. Egami, L. Jiang, T. Nagao, M. Ouchi, M. A. Malkan, T. Hattori, K. Ota, Y. Taniguchi, S. Okamura, C. Ly, M. Iye, H. Furusawa, Y. Shioya, T. Shibuya, Y. Ishizaki, and J. Toshikawa
2011. Completing the Census of Lyα Emitters at the Reionization Epoch. , 734:119.

Kennicutt, Jr., R. C.
1983. The rate of star formation in normal disk galaxies. , 272:54–67.

Kimm, T. and R. Cen
2014. Escape Fraction of Ionizing Photons during Reionization: Effects due to Supernova Feedback and Runaway OB Stars. , 788:121.

Kobayashi, M. A. R., T. Totani, and M. Nagashima
2007. Lyα Emitters in Hierarchical Galaxy Formation. , 670:919–927.

Konno, A., M. Ouchi, K. Nakajima, F. Duval, H. Kusakabe, Y. Ono, and K. Shimasaku
2016. Bright and Faint Ends of Lyα Luminosity Functions at z = 2 Determined by the Subaru Survey: Implications for AGNs, Magnification Bias, and ISM H I Evolution. , 823:20.

Konno, A., M. Ouchi, T. Shibuya, Y. Ono, K. Shimasaku, Y. Taniguchi, T. Nagao, M. A. R. Kobayashi, M. Kajisawa, N. Kashikawa, A. K. Inoue, M. Oguri, H. Furusawa, T. Goto, Y. Harikane, R. Higuchi, Y. Komiyama, H. Kusakabe, S. Miyazaki, K. Nakajima, and S.-Y. Wang
2018. SILVERRUSH. IV. Lyα luminosity functions at z = 5.7 and 6.6 studied with 1300 Lyα emitters on the 14-21 deg^2 sky. , 70:S16.

Koopmans, L., J. Pritchard, G. Mellema, J. Aguirre, K. Ahn, R. Barkana, I. van Bemmel, G. Bernardi, A. Bonaldi, F. Briggs, A. G. de Bruyn, T. C. Chang, E. Chapman, X. Chen, B. Ciardi, P. Dayal, A. Ferrara, A. Fialkov, F. Fiore, K. Ichiki, I. T. Illiev, S. Inoue, V. Jelic, M. Jones, J. Lazio, U. Maio, S. Majumdar, K. J. Mack, A. Mesinger, M. F. Morales, A. Parsons, U. L. Pen, M. Santos, R. Schneider, B. Semelin, R. S. de Souza, R. Subrahmanyan, T. Takeuchi, H. Vedantham, J. Wagg, R. Webster, S. Wyithe, K. K. Datta, and C. Trott
2015. The Cosmic Dawn and Epoch of Reionisation with SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, P. 1.

Kormendy, J.
2013. *Secular Evolution in Disk Galaxies*, P. 1.

Kulas, K. R., A. E. Shapley, J. A. Kollmeier, Z. Zheng, C. C. Steidel, and K. N. Hainline
2011. The Kinematics of Multiple-Peaked Ly-alpha Emission in Star-Forming Galaxies at z˜2-3. *ArXiv e-prints*, 11074367.

Kulkarni, G., L. C. Keating, M. G. Haehnelt, S. E. I. Bosman, E. Puchwein, J. Chardin, and D. Aubert
2019. Large Ly α opacity fluctuations and low CMB τ in models of late reionization with large islands of neutral hydrogen extending to z 5.5. , 485:L24–L28.

Kusakabe, H., K. Shimasaku, M. Ouchi, K. Nakajima, R. Goto, T. Hashimoto, A. Konno, Y. Harikane, J. D. Silverman, and P. L. Capak
2018. The stellar mass, star formation rate and dark matter halo properties of LAEs at z
2. .

Lacey, C. G., C. M. Baugh, C. S. Frenk, A. J. Benson, R. G. Bower, S. Cole, V. Gonzalez-Perez, J. C. Helly, C. D. P. Lagos, and P. D. Mitchell
2016. A unified multiwavelength model of galaxy formation. , 462:3854–3911.

Landy, S. D. and A. S. Szalay
1993. Bias and variance of angular correlation functions. , 412:64–71.

Laursen, P., A. O. Razoumov, and J. Sommer-Larsen
2009. Lyα Radiative Transfer in Cosmological Simulations Using Adaptive Mesh Refinement. , 696:853–869.

Laursen, P. and J. Sommer-Larsen
2007. Lyα Resonant Scattering in Young Galaxies: Predictions from Cosmological Simulations. , 657:L69–L72.

Laursen, P., J. Sommer-Larsen, and A. O. Razoumov
2011a. Intergalactic Transmission and Its Impact on the Lyα Line. , 728:52–+.

Laursen, P., J. Sommer-Larsen, and A. O. Razoumov
2011b. Intergalactic Transmission and Its Impact on the Lyα Line. , 728:52.

Le Delliou, M., C. Lacey, C. M. Baugh, B. Guiderdoni, R. Bacon, H. Courtois, T. Sousbie, and S. L. Morris
2005. The abundance of Lyα emitters in hierarchical models. , 357:L11–L15.

Le Delliou, M., C. G. Lacey, C. M. Baugh, and S. L. Morris
2006. The properties of Lyα emitting galaxies in hierarchical galaxy formation models. , 365:712–726.

Lyth, D. H. and A. R. Liddle
2009. *The Primordial Density Perturbation.*

Malhotra, S. and J. E. Rhoads
2002. Large Equivalent Width Lyα line Emission at z=4.5: Young Galaxies in a Young Universe? , 565:L71–L74.

Matthee, J., D. Sobral, P. Best, I. Smail, F. Bian, B. Darvish, H. Röttgering, and X. Fan
2017. Boötes-HiZELS: an optical to near-infrared survey of emission-line galaxies at z = 0.4-4.7. , 471:629–649.

McQuinn, M., L. Hernquist, M. Zaldarriaga, and S. Dutta
2007. Studying reionization with Lyα emitters. , 381:75–96.

Mitchell, P. D., C. G. Lacey, C. M. Baugh, and S. Cole
2013. How well can we really estimate the stellar masses of galaxies from broad-band photometry? , 435:87–114.

Mo, H., F. C. van den Bosch, and S. White
2010. *Galaxy Formation and Evolution.*

Moresco, M., Euclid Consortium, and Euclid Italia
2018. Euclid: A Space Mission to Study the Origin of the Accelerating Universe. In *VST in the Era of the Large Sky Surveys*, P. 34.

Nagamine, K., M. Ouchi, V. Springel, and L. Hernquist
2010. Lyman-α Emitters and Lyman-Break Galaxies at z = 3-6 in Cosmological SPH Simulations. , 62:1455–.

Nelson, D., A. Pillepich, S. Genel, M. Vogelsberger, V. Springel, P. Torrey, V. Rodriguez-Gomez, D. Sijacki, G. F. Snyder, B. Griffen, F. Marinacci, L. Blecha, L. Sales, D. Xu, and L. Hernquist
2015. The illustris simulation: Public data release. *Astronomy and Computing*, 13:12–37.

Neufeld, D. A.
1990. The transfer of resonance-line radiation in static astrophysical media. , 350:216–241.

Neufeld, D. A.
1991. The escape of Lyman-alpha radiation from a multiphase interstellar medium. , 370:L85–L88.

Orlitová, I., A. Verhamme, A. Henry, C. Scarlata, A. Jaskot, M. S. Oey, and D. Schaerer
2018a. Puzzling Lyman-alpha line profiles in green pea galaxies. , 616:A60.

Orlitová, I., A. Verhamme, A. Henry, C. Scarlata, A. Jaskot, M. S. Oey, and D. Schaerer
2018b. Puzzling Lyman-alpha line profiles in green pea galaxies. , 616:A60.

Orsi, A., C. G. Lacey, and C. M. Baugh
2012. Can galactic outflows explain the properties of Ly α emitters? , 425:87–115.

Orsi, A., C. G. Lacey, C. M. Baugh, and L. Infante
2008. The clustering of Lyα emitters in a ΛCDM Universe. , 391:1589–1604.

Orsi, Á. A. and R. E. Angulo
2018. The impact of galaxy formation on satellite kinematics and redshift-space distortions. , 475:2530–2544.

Osterbrock, D. E.
1989. *Astrophysics of gaseous nebulae and active galactic nuclei.*

Ouchi, M., Y. Harikane, T. Shibuya, K. Shimasaku, Y. Taniguchi, A. Konno, M. Kobayashi, M. Kajisawa, T. Nagao, Y. Ono, A. K. Inoue, M. Umemura, M. Mori, K. Hasegawa, R. Higuchi, Y. Komiyama, Y. Matsuda, K. Nakajima, T. Saito, and S.-Y. Wang
2018. Systematic Identification of LAEs for Visible Exploration and Reionization Research Using Subaru HSC (SILVERRUSH). I. Program strategy and clustering properties of
2000 Lyα emitters at z = 6-7 over the 0.3-0.5 Gpc^2 survey area. , 70:S13.

Ouchi, M., K. Shimasaku, M. Akiyama, K. Sekiguchi, H. Furusawa, S. Okamura, N. Kashikawa, M. Iye, T. Kodama, T. Saito, T. Sasaki, C. Simpson, T. Takata, T. Yamada, H. Yamanoi, M. Yoshida, and M. Yoshida
2005. The Discovery of Primeval Large-Scale Structures with Forming Clusters at Redshift 6. , 620:L1–L4.

Ouchi, M., K. Shimasaku, M. Akiyama, C. Simpson, T. Saito, Y. Ueda, H. Furusawa, K. Sekiguchi, T. Yamada, T. Kodama, N. Kashikawa, S. Okamura, M. Iye, T. Takata,

M. Yoshida, and M. Yoshida
2008. The Subaru/XMM-Newton Deep Survey (SXDS). IV. Evolution of Lyα Emitters from z=3.1 to 5.7 in the 1 deg$^2 Field : LuminosityFunctionsandAGN., 176 : 301 - -330.$

Ouchi, M., K. Shimasaku, H. Furusawa, T. Saito, M. Yoshida, M. Akiyama, Y. Ono, T. Yamada, K. Ota, N. Kashikawa, M. Iye, T. Kodama, S. Okamura, C. Simpson, and M. Yoshida
2010. Statistics of 207 Lyα Emitters at a Redshift Near 7: Constraints on Reionization and Galaxy Formation Models. , 723:869–894.

Oyarzún, G. A., G. A. Blanc, V. González, M. Mateo, and J. I. Bailey, III
2017. A Comprehensive Study of Lyα Emission in the High-redshift Galaxy Population. , 843:133.

Parsa, S., J. S. Dunlop, and R. J. McLure
2018. No evidence for a significant AGN contribution to cosmic hydrogen reionization. , 474:2904–2923.

Partridge, R. B. and P. J. E. Peebles
1967. Are Young Galaxies Visible? , 147:868–+.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay
2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Peebles, P. J. E.
1993. *Principles of physical cosmology.*

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al.
2016. Planck 2015 results. XIII. Cosmological parameters. , 594:A13.

Pritchard, J. R. and A. Loeb
2012. 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901.

Rhoads, J. E., S. Malhotra, A. Dey, D. Stern, H. Spinrad, and B. T. Jannuzi
2000. First Results from the Large-Area Lyman Alpha Survey. , 545:L85–L88.

Santos, M. R., R. S. Ellis, J. Kneib, J. Richard, and K. Kuijken
2004. The Abundance of Low-Luminosity Lyα Emitters at High Redshift. , 606:683–701.

Schaerer, D. and A. Verhamme
2008. 3D Lyα radiation transfer. II. Fitting the Lyman break galaxy MS 1512-cB58 and implications for Lyα emission in high-z starbursts. , 480:369–377.

Shapley, A. E., C. C. Steidel, M. Pettini, and K. L. Adelberger
2003. Rest-Frame Ultraviolet Spectra of z˜3 Lyman Break Galaxies. , 588:65–89.

Shibuya, T., M. Ouchi, Y. Harikane, M. Rauch, Y. Ono, S. Mukae, R. Higuchi, T. Kojima, S. Yuma, C.-H. Lee, H. Furusawa, A. Konno, C. L. Martin, K. Shimasaku, Y. Taniguchi, M. A. R. Kobayashi, M. Kajisawa, T. Nagao, T. Goto, N. Kashikawa, Y. Komiyama, H. Kusakabe, R. Momose, K. Nakajima, M. Tanaka, and S.-Y. Wang
2018. SILVERRUSH. III. Deep optical and near-infrared spectroscopy for Lyα and UV-nebular lines of bright Lyα emitters at z = 6-7. , 70:S15.

Shimakawa, R., T. Kodama, T. Shibuya, N. Kashikawa, I. Tanaka, Y. Matsuda, K.-i. Tadaki, Y. Koyama, M. Hayashi, T. L. Suzuki, and M. Yamamoto
2017. Similarities and uniqueness of Ly α emitters among star-forming galaxies at z = 2.5. , 468:1123–1141.

Sobral, D., J. Matthee, P. Best, A. Stroe, H. Röttgering, I. Oteo, I. Smail, L. Morabito, and A. Paulino-Afonso
2017a. The CALYMHA survey: Lyα luminosity function and global escape fraction of Lyα photons at z = 2.23. , 466:1242–1258.

Sobral, D., J. Matthee, P. Best, A. Stroe, H. Röttgering, I. Oteo, I. Smail, L. Morabito, and A. Paulino-Afonso
2017b. The CALYMHA survey: Lyα luminosity function and global escape fraction of Lyα photons at z = 2.23. , 466:1242–1258.

Sobral, D., J. Matthee, B. Darvish, D. Schaerer, B. Mobasher, H. J. A. Röttgering, S. Santos, and S. Hemmati
2015. Evidence for PopIII-like Stellar Populations in the Most Luminous Lyman-α Emitters at the Epoch of Reionization: Spectroscopic Confirmation. , 808:139.

Sobral, D., J. Matthee, B. Darvish, I. Smail, P. N. Best, L. Alegre, H. Röttgering, B. Mobasher, A. Paulino-Afonso, A. Stroe, and I. Oteo
2018a. The nature of luminous Ly α emitters at z
2-3: maximal dust-poor starbursts and highly ionizing AGN. , 477:2817–2840.

Sobral, D., J. Matthee, B. Darvish, I. Smail, P. N. Best, L. Alegre, H. Röttgering, B. Mobasher, A. Paulino-Afonso, A. Stroe, and I. Oteo
2018b. The nature of luminous Ly α emitters at z
2-3: maximal dust-poor starbursts and highly ionizing AGN. , 477:2817–2840.

Spinoso, D., S. Bonoli, M. Dotti, L. Mayer, P. Madau, and J. Bellovary
2017. Bar-driven evolution and quenching of spiral galaxies in cosmological simulations. , 465:3729–3740.

Spitzer, Jr., L.
1978. Review of Publications: Physical Processes in the Interstellar Medium. , 72:349.

Steidel, C. C., K. L. Adelberger, M. Dickinson, M. Giavalisco, M. Pettini, and M. Kellogg
1998. A Large Structure of Galaxies at Redshift Z approximately 3 and Its Cosmological Implications. , 492:428–438.

Steidel, C. C., M. Bogosavljević, A. E. Shapley, J. A. Kollmeier, N. A. Reddy, D. K. Erb, and M. Pettini
2011. Diffuse Lyα Emitting Halos: A Generic Property of High-redshift Star-forming Galaxies. , 736:160–+.

Steidel, C. C., D. K. Erb, A. E. Shapley, M. Pettini, N. Reddy, M. Bogosavljević, G. C. Rudie, and O. Rakic
2010. The Structure and Kinematics of the Circumgalactic Medium from Far-ultraviolet Spectra of z ˜= 2-3 Galaxies. , 717:289–322.

Steidel, C. C., M. Giavalisco, M. Pettini, M. Dickinson, and K. L. Adelberger
1996. Spectroscopic Confirmation of a Population of Normal Star-forming Galaxies at Redshifts Z 3. , 462:L17+.

Taniguchi, Y., M. Ajiki, T. Nagao, Y. Shioya, T. Murayama, N. Kashikawa, K. Kodaira, N. Kaifu, H. Ando, H. Karoji, M. Akiyama, K. Aoki, M. Doi, S. S. Fujita, H. Furusawa, T. Hayashino, F. Iwamuro, M. Iye, N. Kobayashi, T. Kodama, Y. Komiyama, Y. Matsuda, S. Miyazaki, Y. Mizumoto, T. Morokuma, K. Motohara, K. Nariai, K. Ohta, Y. Ohyama, S. Okamura, M. Ouchi, T. Sasaki, Y. Sato, K. Sekiguchi, K. Shimasaku, H. Tamura, M. Umemura, T. Yamada, N. Yasuda, and M. Yoshida
2005. The SUBARU Deep Field Project: Lymanα Emitters at a Redshift of 6.6. , 57:165–182.

Tepper-García, T.
2006. Voigt profile fitting to quasar absorption lines: an analytic approximation to the Voigt-Hjerting function. , 369:2025–2035.

Verhamme, A., T. Garel, E. Ventou, T. Contini, N. Bouché, E. Herenz, J. Richard, R. Bacon, K. Schmidt, M. Maseda, R. Marino, J. Brinchmann, S. Cantalupo, J. Caruana, B. Clément, C. Diener, A. Drake, T. Hashimoto, H. Inami, J. Kerutt, W. Kollatschny, F. Leclercq, V. Patrício, J. Schaye, L. Wisotzki, and J. Zabl
2018. Recovering the systemic redshift of galaxies from their Lyman alpha line profile. , 478:L60–L65.

Verhamme, A., D. Schaerer, H. Atek, and C. Tapken
2008. 3D Lyα radiation transfer. III. Constraints on gas and stellar properties of z ˜ 3 Lyman break galaxies (LBG) and implications for high-z LBGs and Lyα emitters. , 491:89–111.

Verhamme, A., D. Schaerer, and A. Maselli
2006. 3D Lyα radiation transfer. I. Understanding Lyα line profile morphologies. , 460:397–413.

Vilella-Rojo, G., K. Viironen, C. López-Sanjuan, A. J. Cenarro, J. Varela, L. A. Díaz-García, D. Cristóbal-Hornillos, A. Ederoclite, A. Marín-Franch, and M. Moles
2015. Extracting Hα flux from photometric data in the J-PLUS survey. , 580:A47.

Wang, Y., M. Robberto, M. Dickinson, H. C. Ferguson, L. A. Hillenbrand, W. Fraser, P. Behroozi, J. Brinchmann, A. Cimatti, R. Content, E. Daddi, C. Hirata, M. J. Hudson, J. D. Kirkpatrick, A. Orsi, M. Ballardini, R. Barkhouser, J. Bartlett, R. Benjamin, R. Chary, C.-H. Chuang, C. Conroy, M. Donahue, O. Dore, P. Eisenhardt,

K. Glazebrook, G. Helou, S. Malhotra, L. Moscardini, J. A. Newman, Z. Ninkov, M. Ressler, J. Rhoads, J. Rhodes, D. Scolnic, A. Shapley, S. Smee, F. Valentino, and R. H. Wechsler
2018a. ATLAS Probe: Breakthrough Science of Galaxy Evolution, Cosmology, Milky Way, and the Solar System. *ArXiv e-prints.*

Wang, Y., M. Robberto, M. Dickinson, L. A. Hillenbrand, W. Fraser, P. Behroozi, J. Brinchmann, C.-H. Chuang, A. Cimatti, R. Content, E. Daddi, H. C. Ferguson, C. Hirata, M. J. Hudson, J. D. Kirkpatrick, A. Orsi, R. Ryan, A. Shapley, M. Ballardini, R. Barkhouser, J. Bartlett, R. Benjamin, R. Chary, C. Conroy, M. Donahue, O. Dore, P. Eisenhardt, K. Glazebrook, G. Helou, S. Malhotra, L. Moscardini, J. A. Newman, Z. Ninkov, M. Ressler, J. Rhoads, J. Rhodes, D. Scolnic, S. Smee, F. Valentino, and R. H. Wechsler
2018b. ATLAS Probe: Breakthrough Science of Galaxy Evolution, Cosmology, Milky Way, and the Solar System. *arXiv e-prints.*

Weinberger, L. H., M. G. Haehnelt, and G. Kulkarni
2019. Modelling the observed luminosity function and clustering evolution of Ly α emitters: growing evidence for late reionization. , 485:1350–1366.

Weinberger, L. H., G. Kulkarni, M. G. Haehnelt, T. R. Choudhury, and E. Puchwein
2018. Lyman-α emitters gone missing: the different evolution of the bright and faint populations. , 479:2564–2587.

White, H. E.
1934. *Introduction to atomic spectra.*

Yang, H., S. Malhotra, J. E. Rhoads, and J. Wang
2017. Blueberry Galaxies: The Lowest Mass Young Starbursts. , 847:38.

Zheng, Z., R. Cen, H. Trac, and J. Miralda-Escudé
2010. Radiative Transfer Modeling of Lyα Emitters. I. Statistics of Spectra and Luminosity. , 716:574–598.

Zheng, Z., R. Cen, H. Trac, and J. Miralda-Escudé
2011. Radiative Transfer Modeling of Lyα Emitters. II. New Effects on Galaxy Clustering. , 726:38–+.

Zheng, Z. and J. Miralda-Escudé
2002. Monte Carlo Simulation of Lyα Scattering and Application to Damped Lyα Systems. , 578:33–42.

Zheng, Z. and J. Wallace
2014. Anisotropic Lyman-alpha Emission. , 794:116.

Thanks for reading!